DOCUMENT RESUME

ED 061 631 EA 004 236

AUTHOR	Ray, H. W.; And Others
TITLE	The Office of Economic Opportunity Experiment in Educational Performance Contracting. Final Report.
INSTITUTION	Battelle Memorial Inst., Columbus, Ohio. Columbus Labs.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
PUB DATE	14 Mar 72
NOTE	298p.; For additional information, see EA 004 259 (Battelle Interim Report) and EA 004 260 (Education Turnkey Systems Final Report)
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	*Academic Achievement; *Achievement Gains; Educational Change; *Educational Experiments; *Educational Research; Improvement Programs; Individualized Instruction; Mathematics Instruction; Measurement; *Performance Contracts; Reading Instruction; Skill Development; Spanish Speaking; Standardized Tests; Teaching Methods; Testing; Test Results

ABSTRACT

This report represents an integration of analysis results in an interim report with those from additional analyses performed since. It is divided into (1) description of the experiment in which the goal is outlined; (2) description of the technology company programs, an overview of personnel, curriculum and materials, and incentive systems; (3) description of the target population which provides variables of socioeconomic status; (4) criteria employed for the selection of achievement tests used and a description of tests selected; (5) design and procedures used to administer selected tests and a summary of conditions for both pre- and post testing; (6) the data-analysis method used for the analysis of student achievement test data and the rationale for the analysis method employed; (7) results and conclusions: technology company sites; and (8) results and conclusions: incentive only sites. Findings reveal little evidence that the performance contracting experiments at technology company sites or those at incentives only sites had beneficial effects on the reading or mathematics achievement of participating students as measured by a standardized achievement test. Several appendixes provide basic statistical data for readers to examine and analyze. A related document is EA 004 144. (Author/JF)

ED 061631

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

RESEARCH REPORT

Battelle
Columbus Laboratories

EA 004 236

FINAL REPORT

on

THE OFFICE OF ECONOMIC OPPORTUNITY EXPERIMENT IN EDUCATIONAL PERFORMANCE CONTRACTING

to

OFFICE OF ECONOMIC OPPORTUNITY

March 14, 1972

Project Director: H. W. Ray

Project Team:

Clifford Audretch
Richard W. Clark
Roger W. Cote
Kenneth Eckhardt
Dennis N. McFadden
Michael P. Penn
John R. Powers III
E. Allen Schenck
John R. Stock
Ralph E. Thomas
Mary Beth Zak

BATTELLE
Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

TABLE OF CONTENTS

Page

INTRODUCTION 1

DESCRIPTION OF THE EXPERIMENT. 3

- Selection of School Districts 5
- Selection of Technology Companies 6
- Selection of Schools and the Creation of Experimental and Control Groups. 7
- Selection of Students 9
- Experimental Procedures 10
- The Incentive Only School Districts 12

DESCRIPTION OF TECHNOLOGY COMPANY PROGRAMS 14

- Alpha Learning Systems, Incorporated 14
- Learning Foundations, Incorporated. 15
- Plan Education Centers, Incorporated 16
- Quality Educational Development, Incorporated 17
- Singer/Graflex. 18
- Westinghouse Learning Centers 19
- Tabular Summary 21

DESCRIPTION OF THE TARGET POPULATION 23

- Securing Family Background Information. 24
- Characteristics of Participants 25
- Comparison of Full-Year Students and Dropouts 39

ACHIEVEMENT MEASURES 43

- Selection of Evaluation Tests 43
- Descriptions of Evaluation Tests Selected 46
- Certification Tests 49

TESTING PROCEDURES AND CONDITIONS. 53

- Pretesting. 53
- Posttesting 62

THE DATA ANALYSIS METHOD 67

- Issues in the Selection of an Analysis Method 67
- Selection of an Analysis Method 73
- Regression Analysis Models Employed 75

TABLE OF CONTENTS
(Continued)

Page

RESULTS AND CONCLUSIONS: TECHNOLOGY COMPANY SITES 84

Introduction. 84

The Main Analyses: Regression Analyses of Experimental versus Control Groups on First Year Test Results. . . . 86

Experimental - Control Group Comparisons Aggregated Across Sites. 105

Experimental Versus Control Groups: Second-Year Stability of Impact Results 109

Comparison Groups Analysis. 116

Analysis of Special Treatment Groups. 123

Analysis of Attendance Data 136

Summary of Results and Conclusions. 140

RESULTS AND CONCLUSIONS: INCENTIVES ONLY SITES. 143

Experimental Versus Control Groups: Regression Analyses of First-Year Results. 143

Stability of Impact Results. 148

Analysis of Attendance Data 148

Summary of Results and Conclusions. 150

APPENDIX A

NUMBER OF FULL-YEAR STUDENTS BY SITE, GROUP, AND GRADE . . . A-1

APPENDIX B

SURVEY RETURN RATE BY SITE, GROUP, AND GRADE B-1

APPENDIX C

PARENT QUESTIONNAIRE C-1

APPENDIX D

PERCENTAGE DISTRIBUTION OF FULL-YEAR STUDENTS' RACE, BY SITE AND GROUP D-1

APPENDIX E

MEDIAN TOTAL FAMILY INCOME BY SITE, GROUP, AND GRADE E-1

TABLE OF CONTENTS
(Continued)

Page

APPENDIX F

PERCENTAGE DISTRIBUTION OF FATHER'S EDUCATION, FOR FULL-YEAR STUDENTS, BY SITE AND GROUP. F-1

APPENDIX G

PERCENTAGE DISTRIBUTION OF FATHER'S OCCUPATION FOR FULL-YEAR STUDENTS, BY SITE AND GROUP. G-1

APPENDIX H

ENTRY LEVEL ACHIEVEMENT DATA FOR FULL-YEAR STUDENTS H-1

APPENDIX I

PERCENTAGE DISTRIBUTION OF RESPONSES TO ITEM NO. 10 (APPROVE NEW METHODS) OF PARENT QUESTIONNAIRE. I-1

APPENDIX J

DESCRIPTION OF CERTIFICATION TESTS. J-1

APPENDIX K

MEANS, STANDARD DEVIATIONS, AND CORRELATIONS BETWEEN PRE- AND POSTTEST SCORES IN READING AND MATHEMATICS FOR FULL-YEAR STUDENTS WITH BOTH PRE- AND POSTTEST SCORES BY SITE AND GROUP K-1

APPENDIX L

ILLUSTRATIVE SCATTERPLOTS L-1

APPENDIX M

PRE-POST REGRESSIONS HAVING SIGNIFICANT GROUP-BY-PRETEST INTERACTIONS M-1

APPENDIX N

SAMPLE SIZE FOR PRE-POST MODEL REGRESSION ANALYSES. N-1

TABLE OF CONTENTS
(Continued)

Page

APPENDIX O

MEANS AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SCORES FOR EXPERIMENTAL AND COMPARISON STUDENTS, BY SITE, GRADE AND SUBJECT AREA O-1

APPENDIX P

MEANS AND STANDARD DEVIATIONS OF PRETEST AND SECOND YEAR TEST SCORES FOR FULL-YEAR EXPERIMENTAL AND CONTROL STUDENTS HAVING BOTH OF THESE SCORES, BY SITE, GRADE, AND GROUP . . P-1

APPENDIX Q

MEANS AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SCORES FOR FULL-YEAR EXPERIMENTAL AND SPECIAL TREATMENT STUDENTS BY PROGRAM, SITE, GRADE, AND SUBJECT AREA Q-1

APPENDIX R

ESTIMATED POSTTEST MEANS FOR EXPERIMENTAL (E) AND CONTROL (C) GROUPS, DIFFERENCES BETWEEN THESE MEANS, AND ASSOCIATED t RATIOS, FOR ALL SITES, GRADES, AND SUBJECTS. R-1

APPENDIX S

REGRESSION RESULTS FOR SECOND-YEAR TESTING S-1

APPENDIX T

REGRESSION RESULTS FOR CONTROL VERSUS COMPARISON GROUPS T-1

APPENDIX U

REGRESSION RESULTS FOR EXPERIMENTAL VERSUS COMPARISON GROUPS U-1

APPENDIX V

REGRESSION RESULTS FOR SPECIAL TREATMENT GROUPS AT GRAND RAPIDS V-1

APPENDIX W

REGRESSION RESULTS FOR SPECIAL TREATMENT GROUPS AT HARTFORD . W-1

APPENDIX X

REGRESSION RESULTS FOR ATTENDANCE X-1

LIST OF TABLES

Page

Table 1. Comparison of Particular Aspects of Experimental Programs. 22

Table 2. Predominant Races of Full-Year Students, By Site and Group 28

Table 3. Median Total Family Income and Percentage of Families Having and Income Less than the National Average For Full-Year Students, by Site and Group 29

Table 4. Percentage Distribution of Father's Education for Full-Year Students, by Group, for All Sites . 32

Table 5. Percentage Distribution of Father's Occupation for Full-Year Students, by Group, for All Sites . 32

Table 6. Percentage of Full-Year Students Below Grade Level for Reading and Mathematics, by Group and Grade, for All Sites 35

Table 7. Overall Mean Pretest Grade Equivalent Values for Reading and Mathematics, by Group and Grade, for All Sites 37

Table 8. Additional Variables, Other Than Pretest, on Which the Experimental and Control Groups Were Judged to Differ for Each Site 40

Table 9. Mean Pretest Values (and Associated Grade Equivalents) for Full-Year Experimental Students and Dropouts, by Grade and Subject Area 41

Table 10. List of Tests for Each Grade. 45

Table 11. Reliability Coefficients and Number of Items for the First Grade Pretest (Stanford Early Achievement). 50

Table 12. Reliability Coefficients and Number of Items for the First Grade Posttest (California Achievement Test) 50

Table 13. Reliability Coefficients and Number of Test Items for the Second Grade Test (MET 70, Primary I) . . 51

Table 14. Reliability Coefficients and Number of Test Items for the Third Grade Test (MET 70, Primary II) . . 51

LIST OF TABLES
(Continued)

Page

Table 15. Reliability Coefficients and Number of Test Items for the Seventh Grade (MET 70, Intermediate). . . 52

Table 16. Reliability Coefficients and Number of Test Items for the Eighth and Ninth Grade Test (MET 70, Advanced) 52

Table 17. Testing Incidents Reported for Pre-Evaluation Testing 58

Table 18. Testing Incidents Reported for Post-Evaluation Testing 65

Table 19. Estimated Reliability Coefficients (Kuder-Richardson 21) of the Pre- and Posttest Forms of the Evaluation Reading and Mathematics Tests for Each Grade. 71

Table 20. Distribution of Pre-Post Correlations for Experimental and Control Groups, by Grade and Subject Area 73

Table 21. Estimated Experimental (E) and Control (C) Posttest Means in Reading and Mathematics and Differences Between These Means for Sites Exhibiting a Significant Group Difference in Grade 1 89

Table 22. Estimated Experimental (E) and Control (C) Posttest Means in Reading and Mathematics and Differences Between These Means for Sites Exhibiting a Significant Group Difference in Grade 2 91

Table 23. Estimated Experimental (E) and Control (C) Posttest Means in Reading and Mathematics and Differences Between These Means for Sites Exhibiting a Significant Group Difference in Grade 3 92

Table 24. Estimated Experimental (E) and Control (C) Posttest Means in Reading and Mathematics and Differences Between These Means for Sites Exhibiting a Significant Group Difference in Grade 7 94

LIST OF TABLES
(Continued)

Page

Table 25. Estimated Experimental (E) and Control (C) Posttest Means in Reading and Mathematics and Differences Between These Means for Sites Exhibiting a Significant Group Difference in Grade 8. 95

Table 26. Estimated Experimental (E) and Control (C) Posttest Means in Reading and Mathematics and Differences Between These Means for Sites Exhibiting a Significant Group Difference in Grade 9. 96

Table 27. Summary of Pre-Post Model Results for Each Grade, Site, and Subject 98

Table 28. Number of Positive Impacts From the Pre-Post Model Analyses for the Elementary and Junior High Grades, for Each Technology Company 100

Table 29. Independent Variables Constituting the Extended Variables Model for Each Grade/Site Combination. 102

Table 30. Comparison of Significant E/C Group Differences Found Using the Pre-Post Model (PPM) Versus the Extended Variables Model (EVM) for Each Grade, Site, and Subject. 103

Table 31. Overall Mean Pretest and Posttest Values (and Associated Grade Equivalents) for Experimental (E) and Control (C) Students, by Subject Area and Grade. 107

Table 32. Site/Grade/Subject Combinations Where Second-Year Testing Was Accomplished. 110

Table 33. Estimated Experimental (E) and Control (C) Second-Year Test Means for Sites Showing a Significant Group Difference 112

Table 34. Statistical Outcomes of the Pre-Post Model Analysis of Spring (1971) and Fall (1971) Test Performance for All Grade/Site/Subjects for Which There Was Second-Year Testing. 113

Table 35. Classification of Results to Show Stability of Program Impact 114

LIST OF TABLES
(Continued)

Page

Table 36. Sites and Grades Where Comparison Groups Were Posttested 116

Table 37. Summary of Significant Differences Between Control (C) Groups and Comparison (R) Groups . . . 118

Table 38. Summary of Significant Differences Between Experimental (E) and Comparison (R) Groups 119

Table 39. Paired Results Obtained From Experimental Versus Control Comparisons and Comparison Versus Control Comparisons. 121

Table 40. Special Programs at Grand Rapids and Hartford. . . 124

Table 41. Estimated Posttest Means for Control (C) and Special Treatment (ST) Groups for Groups Showing Significant Differences at Grand Rapids, Michigan 127

Table 42. Estimated Posttest Means for Control (C) and Special Treatment (ST) Groups for Groups Showing Significant Differences at Hartford, Connecticut. 128

Table 43. Estimated Posttest Means for Experimental (E) and Special Treatment (ST) Groups for Groups Showing Significant Differences at Grand Rapids, Michigan 130

Table 44. Estimated Posttest Means for Experimental (E) and Special Treatment (ST) Groups for Groups Showing Significant Differences at Hartford, Connecticut. 131

Table 45. Summary of Results for the Special Treatment Groups. 133

Table 46. Sites and Grades Where Attendance Analyses Were Performed. 136

Table 47. Estimated Mean 1970-71 Percentage of School Days Absent for Experimental (E) and Control (C) Groups Showing Significant Differences 138

LIST OF TABLES
(Continued)

Page

Table 48. Estimated Experimental (E) and Control (C) Posttest Means and Differences Between These Means for Those Grade/Site/Subject Combinations Exhibiting a Significant Group Difference. 145

Table 49. Summary of Pre-Post Model Results for Each Grade and Subject at Mesa and Stockton. 146

Table 50. Independent Variables Constituting the Extended Variables Model for Each Grade/Site Combination. 147

Table 51. Comparison of Significant E/C Group Differences Found Using the Pre-Post Model Versus the Extended Variables Model for Each Grade, Site and Subject. 148

Table 52. Regression Results for Second Year-Testing at MESA. 149

Table 53. Regression Results for Attendance at MESA. 149

Table A-1. Number of Full-Year Students by Site, Group, and Grade A-1

Table B-1. Survey Return Rate by Site, Group, and Grade. . . B-1

Table D-1. Percentage Distribution of Full-Year Students' Race, by Site and Group D-1

Table E-1. Median Total Family Income by Site, Group, and Grade E-1

Table F-1. Percentage Distribution of Father's Education, For Full-Year Students, by Site and Group F-1

Table G-1. Percentage Distribution of Father's Occupation, For Full-Year Students, by Site and Group G-1

Table H-1. Mean GEQ Values of Full-Year Students for Reading, by Site, Group, and Grade H-1

Table H-2. Mean GEQ Values of Full-Year Students for Math, by Site, Group, and Grade H-2

Table H-3. Percentage of Full-Year Students Below Grade Level for Reading, by Site, Group, and Grade. . . H-3

Table H-4. Percentage of Full-Year Students Below Grade Level for Mathematics, by Site, Group, and Grade H-4

Table I-1. Percentage Distribution of Responses to Item No. 10 (Approve New Method) of Parents of Full-Year Students by Site and Group. I-1

LIST OF TABLES
(Continued)

Page

Table J-1. Listing of Certification Tests. J-1

Table J-2. Reliability Coefficients, Standard Errors of Measurement, and Number of Items for the Third Grade Certification Tests J-9

Table J-3. Reliability Coefficients, Standard Errors of Measurement, and Number of Items for a National Sample, CAT, Level 4, Grade 7 . . . . J-11

Table J-4. Reliability Coefficients, Standard Errors of Measurement, and Number of Items for a National Sample, CTBS, Level 3, Grade 6. . . . . J-13

Table J-5. Reliability Coefficients, Standard Errors of Measurement, and Number of Items for the ITBS, Grade 8 . . . . J-14

Table K-1. Means, Standard Deviations and Correlations Between Pre- and Posttest Scores in Reading for Full-Year Students With Both Pre and Posttest Scores by Site and Group K-1

Table K-2. Means, Standard Deviations and Correlations Between Pre- and Posttest Scores in Mathematics for Full-Year Students With Both Pre- and Posttest Scores by Site and Group K-7

Table M-1. Definitions of Types of Regression Plots for Posttest Versus Pretest for E and C Groups . . . M-2

Table M-2. Classification of Pre-Post Regression Plots Having Significant Group-By-Pretest Interactions M-4

Table N-1. Number of Experimental (E), Control (c), and Comparison (R) Full-Year Students Having Both A Pretest and Posttest Score, by Site, Grade, and Subject Area. N-1

Table N-2. Number of Full-Year Experimental (E) and Control (C) Students Given a Retention Test Having Both a Pretest and Retention Test Score, by Site, Grade, and Subject Area N-4

Table N-3. Number of Experimental (E) and Control (C) Full-Year Students Having Both 1969-70 and 1970-71 Attendance Data, by Site and Grade. N-5

LIST OF TABLES
(Continued)

Page

Table N-4. Number of Special Treatment Full-Year Students Having Both a Pretest and Posttest Score, by Program, Site, Grade, and Subject Area. N-6

Table O-1. Means (and Associated Grade Equivalents) and Standard Deviations of Pretest and Posttest Scores for Full-Year Experimental and Comparison Students Having Both of These Scores, by Site, Grade, and Subject Area O-1

Table P-1. Means (and Associated Grade Equivalents) and Standard Deviations of Pretest and Second-Year Test Scores for Full-Year Experimental and Control Students Having Both of These Scores, by Site, Grade, Subject, and Group P-1

Table Q-1. Means (and Associated Grade Equivalents) and Standard Deviations of Pretest and Posttest Scores for Full-Year Experimental and Special Treatment Students Having Both of These Scores by Program, Site, Grade, and Subject Area Q-1

Table R-1. Pre-Post Model Results R-1

Table S-1. Regression Results for Second-Year Testing. . . . S-1

Table T-1. Regression Results for Control (C) Versus Comparison (R) Groups T-1

Table U-1. Regression Results for Experimental (E) Versus Comparison (R) Groups U-1

Table V-1. Regression Results in Reading for Experimental (E) Versus Special Treatment (ST) Groups at Grand Rapids. V-1

Table V-2. Regression Results in Mathematics for Experimental (E) Versus Special Treatment (ST) Groups at Grand Rapids. V-3

Table V-3. Regression Results in Reading for Control (C) Versus Special Treatment (ST) Groups at Grand Rapids. V-5

Table V-4. Regression Results in Mathematics for Control (C) Versus Special Treatment (ST) Groups at Grand Rapids. V-7

LIST OF TABLES
(Continued)

Page

Table W-1. Regression Results in Reading for Experimental (E) Versus Special Treatment (ST) Groups at Hartford. W-1

Table W-2. Regression Results in Mathematics for Experimental (E) Versus Special Treatment (ST) Groups at Hartford. W-2

Table W-3. Regression Results in Reading for Control (C) Versus Special Treatment (ST) Groups at Hartford. W-3

Table W-4. Regression Results in Mathematics for Control (C) Versus Special Treatment (ST) Groups at Hartford. W-4

Table X-1. Regression Results for Attendance X-1

LIST OF FIGURES

Figure 1. Overall Average Pretest Grade Equivalents for Experimental and Control Groups, for Each Grade in Reading and Mathematics 36

Figure 2. Illustration of the "Pre-Post Model". 77

Figure 3. Grade Equivalent Gains by Grade for Mathematics and Reading 106

INTRODUCTION

The purpose of this final report is to present Battelle Columbus Laboratories' principal findings for the OEO Experiment in Educational Performance Contracting. This report represents an integration of analysis results presented in Battelle's Interim Report* with results from additional analyses performed since that time.

This report is divided into eight sections:

(1) Description of The Experiment
(2) Description of The Technology Company Programs
(3) Description of the Target Population
(4) Achievement Measures
(5) Testing Procedures and Conditions
(6) The Data-Analysis Method
(7) Results and Conclusions: Technology Company Sites
(8) Results and Conclusions: Incentive Only Sites.

In section one (Description of Experiment), the goal of the experiment is outlined, along with a summarization of the more important characteristics of the experimental design and experimental procedures.

Section two provides an overview of the technology company programs, in terms of personnel utilized by the companies, curriculum and materials, and the incentive systems and strategies used.

Section three presents descriptive information on some of the more important variables characterizing the socioeconomic status, social and home background, and achievement levels of students who participated in the experiment, so as to afford the reader a description of the type of student for which the results and conclusions of this study apply.

In section four (Achievement Measures), a summary is given of the criteria employed for the selection of achievement tests used to assess program impact on student learning, along with a brief description of the tests selected. Section five, then, outlines the design and procedures used to adminster the selected tests, followed by a summary of the testing conditions for both the pre- and posttesting phases.

---

* Interim Report on The Office of Economic Opportunity Experiment in Educational Performance Contracting, January 29, 1972. Battelle, Columbus Laboratories, 505 King Avenue, Columbus, Ohio, 43201.

Section six outlines the method used for the analysis of student achievement test data, and the rationale for the analysis method employed. This section will be of most benefit to a reader with basic training in statistical methodology. However, section six will also provide the non-statistically trained reader with a general orientation to the analysis methodology employed and the rationale for the use of this method.

Section seven of the report presents the principal results and conclusions concerning the impact of educational performance contracting on student achievement in the eighteen school districts that contracted with private educational technology companies. In the last section of the report, results and conclusions are presented for the two "Incentives Only" sites (Stockton, California, and Mesa, Arizona), where performance incentive contracting was implemented without the use of a private educational technology company.

Sections one, two, four, and five as outlined above contain essentially the same contents as were in Battelle's Interim Report. Section three (Description of the Target Population) is expanded, providing additional information on experimental and control groups of students and differences in the characteristics of these two groups. Sections six, seven, and eight are changed, reflecting additional tasks performed since the Interim Report.

Several Appendices to the report are provided that contain basic statistical data, so as to afford the interested reader an opportunity to examine and analyze detailed data not always presented in the body of the report, but which serve as a basis for the results and conclusions presented.

DESCRIPTION OF THE EXPERIMENT

In view of the inability of traditional educational methods to narrow the gap between the academic performance of low-income, underachieving students and "average" students, and the increasing public concern for accountability in education, OEO implemented an experimental remedial education program in reading and mathematics during the school year 1970-1971. This program involved federal support to participating school districts for the subcontracting of remedial teaching in reading and mathematics to private educational technology companies. These private companies subcontracted with the schools on a performance incentive basis. That is, the companies were to be remunerated for their services in proportion to the achievement gains of students throughout the year.

Contracts were signed with two organizations to provide management support for the implementation of these subcontracts, to provide economic analyses of the costs of the programs, and to provide an evaluation of the educational achievement benefits obtained from each of the programs. Battelle has performed the evaluation of the educational achievement in the experiment. This section of the report includes a description of all facets of the experiment in performance-incentive remedial education which are relevant to the analysis and evaluation of its educational outcomes, based upon testing during school years 1970-71 and 1971-72.

The general goal of this experiment was to determine how effective the innovative programs would be in producing significant gains in the reading and mathematics skills of low-achieving students from low-income families. A general hypothesis that could be stated concerning the experimental programs was that low income, low-achieving students receiving instruction in the experimental programs for 1 year in reading and mathematics would show superior achievement measured by standardized tests to that of similar students receiving normal school instruction in the same subjects.

This experiment can be characterized as an effort to apply the concepts and methodology of laboratory experimentation in a real-world setting, in order to test the research hypothesis. Clearly, the degree of experimental "control" possible in a real-world setting falls far short of that obtainable in a pure laboratory setting. On the other hand, the experiment represents much more than the unstructured collection of observational data in order to answer the main research question. Indeed, considerable efforts were extended to achieve as much structure and control as is possible in a real-world environment, in order to provide a valid test of the research hypothesis.

As with most field experiments, compromises had to be made in the experimental design which made the testing of the research hypothesis difficult. The most notable was the nonrandom assignment of students into the experimental and regular school classes. Random assignment of students was impossible, given the size of the experiment. However, efforts were made to select students for experimental and for regular instruction in a way so as to achieve a "match" between the two groups of students on several important variables affecting student performance. Thus, efforts were made to achieve comparable groups with respect to entry level achievement of the students, and with respect to family background variables reflecting the socioeconomic status of students (e.g., family income level). Further, the design sought to measure several of the pertinent variables affecting student performance, in order to determine the comparability of students receiving experimental and regular instruction, and for the purpose of taking into account in the analysis group differences that might exist.

In order to accomplish the overall goal of the experiment, the following experimental features were adopted by OEO:

(1) A large number of school districts, varying in size, geographical location, rural-urban setting, and minority group composition were selected to participate, in order to provide a comprehensive test of the experimental programs.

(2) Programs were implemented in the same grades in all school districts for the same number of students. Grades 1, 2, 3 and 7, 8, and 9 were selected with enrollment generally 100 students per grade. In total, some 13,000 students participated in experimental programs and 12,000 in control programs.

(3) The performance of students in the experimental programs was compared with that of "control" students within the same school district receiving normal school instructional programs.

(4) Experimental and control schools were selected within districts with the objective of obtaining schools whose student populations were matched on achievement deficiency, family income, and minority group composition.

(5) Sufficient numbers of students were to be selected to enable the study of student characteristics which might be related to student performance.

(6) Technology companies (subcontractors) were selected to represent a variety of educational technology and approaches.

Selection of School Districts

From a pool of over 200 school districts that had expressed interest in the OEO performance-incentive remedial education experiment, 163 replied to an OEO invitation to participate in the experiment. Of these, 77 made formal application to participate. In order to select school districts, OEO applied the following criteria:

(1) Students had to come from "low-income" families, according to ESEA Title I or OEO criteria.

(2) There had to be at least 100 students in each of grades 1,2,3,7,8, and 9 deficient in reading and mathematics achievement who could serve as an experimental population and an equal number to serve as a control population.

(3) There were recent (1968-69 or 1969-70) valid and reliable achievement test data available to facilitate the assessment of student populations within the schools of the district who could serve as experimental and control groups.

(4) The school district had some familiarity with performance contracting.

(5) The school district showed no evidence of having political, social, or economic problems that might interfere with the implementation of the experiment.

Of the 77 school districts that applied, 18 were selected for participation, as listed below.

- Anchorage, Alaska
- Athens, Georgia
- Bronx, New York
- Dallas, Texas
- Fresno, California
- Grand Rapids, Michigan
- Hammond, Indiana
- Hartford, Connecticut
- Jacksonville, Florida
- Las Vegas, Nevada
- McComb, Mississippi
- Philadelphia, Pennsylvania
- Portland, Maine
- Rockland, Maine
- Seattle, Washington
- Selmer, Tennessee
- Taft, Texas
- Wichita, Kansas

The districts were selected to be representative of the many different types that contain low-income, low-achieving populations in the United States. As can be seen the 18 districts included 4 large urban school systems (Bronx, Philadelphia, Dallas, and Seattle), plus several middle-size urban systems. Smaller and rural systems were represented by Athens, McComb, Rockland, Selmer, and Taft.

Selection of Technology Companies

Thirty-one technology companies applied to participate in the experiment in response to a competitive request for proposals issued by OEO. Six firms were selected. The instructional approaches of these firms varied by relative emphasis upon hardware, incentives, or curricular software and teacher training methods. The firms were

(1) Alpha Learning Systems, Inc.
(2) Singer/Graflex, Inc.
(3) Westinghouse Learning Corporation
(4) Quality Educational Development, Inc.
(5) Learning Foundations, Inc.
(6) Plan Education Centers, Inc.

An attempt was made to have each company instruct as many different subpopulations of students (e.g., white, black, Spanish-speaking, etc.) as possible. Consequently, each company was assigned to work in widely varying districts. With 18 districts in total, each of the 6 companies was assigned to work in 3 districts.

These 6 companies provided a range of educational approaches. Their diversity is described elsewhere in this report (Reference Section on Description of Experimental Programs).

Selection of Schools and the Creation of Experimental and Control Groups

Within each selected school district, an effort was made to identify schools which could provide 100 students in each of six grades (1,2,3,7,8 and 9) for the experimental groups. In three of the districts, this requirement was reduced to 75 per grade, so as to include smaller districts in the study. Students had to be deficient in reading and mathematics skills, from low-income families, and be representative of minority groups within the district. Usually at the primary level, more than two schools had to be selected to provide the 100 students in each of grades 1, 2 and 3. This was because of the normal small size of such schools. At the secondary level, one school frequently could be selected, although in some cases more than one was necessary to provide the designated experimental populations.

The basic design called for a comparison of experimental (E) students receiving the innovative instruction with control (C) students receiving normal instruction in other schools within the same districts. The control students were designated in different school buildings in order to prevent any confounding of E versus C comparisons by having the effect of the experimental programs "rub off" or become assimilated through adjacent classrooms. Control schools, too, had to have a population of 100 low-achieving students from low-income families in each of grades 1, 2,3 and 7,8 and 9.

In order to measure any "rub off" effect in the experimental schools, an additional 50 students in each of the six grades were designated as "comparison" students. These students were tested with the experimental students and were also to serve as a replacement pool for students who might leave the experimental programs during the year.

A fourth group was also established in some districts. Where the district already was operating another special reading and math program, OEO offered to test those students and compare them with other groups in the experiment. These were known as "Special Treatment" groups. They were found in Grand Rapids, where two other non-OEO performance contracts and one special school project were operated, and in Hartford.

Since the comparison of the performance of E and C students was of primary interest, the selection of individual schools within a district focused on these populations. Criteria employed for the selection of E and C schools were

(1) The schools had to be ESEA Title I designees, or, if not, at least 80% of the students to be in the project had to meet at least the Title I criteria.

(2) The schools had to be among the most academically deficient Title I schools in the district.

(3) There could be no other "special programs" in reading and mathematics within the particular schools that could affect the progress of groups in the experiment.

(4) As stated, E schools had to contain 100 academically deficient students per grade (75 students per grade in three predominantly rural districts) and at least one-half again as many academically deficient students per grade who would be available as replacement students (the comparison group). In C schools, 100 academically deficient students per grade (75 students per grade in three predominantly rural districts) had to be present.

The process of selecting schools as experimental and control within a district was based upon an analysis of their relative deficiencies in reading and math as shown by the districts' testing data. These were the data required to be present for a district to be considered for inclusion in the experiment. During the summer, a subcontractor to the management support contractor arrayed the district test data to show the relative ranking of Title I eligible schools in reading and math achievement.

Generally speaking, the most deficient school or schools were selected as the experimental school(s) and the next most deficient as the control schools. In large districts where there was sometimes a substantial choice of schools, the selection was affected by the presence or absence of other special programs in them, the receptivity of individual schools to being included in the experiment and the relative efficiency with which the required number of experimental and control students could be accumulated in them. For two of the smaller districts, Rockland and Taft, the control schools had to be selected in an adjoining district.

Selection of Students

Once experimental and control schools were selected, students within these schools were selected into appropriate groups in a similar manner. The management support contractor arrayed reading and math scores of individual students by school to show the lowest achieving first, the next lowest achieving, etc. The student scores were shown in terms of decrements below grade level, i.e., the number of grade levels (in tenths) each student was behind his expected grade placement at the time of testing, as shown by previous years' testing done by schools. Within each experimental and control school, the lowest achieving students were selected first, in order, up to the required number of students per school. In those cases where previous scores in both reading and math were available for a student, an average of the two scores was taken to establish one decrement upon which to rank the students.

In the case of comparison students within the experimental schools, the next 50 lowest achieving students were selected after the 100 experimental students had been selected. Membership in the special treatment groups was determined by student designation through other means into the special programs.

It was not feasible to follow the basic student selection process described above in every case. The basic lists of students for selection were made up during the summer preceding the school year, based upon already existing test data. Several changes usually occurred in each grade in districts because some students moved from the school. In such cases, replacements were designated locally after school had begun. In other cases, if previous year's test data were not available, teacher recommendations about low-achieving students who could benefit from the program were used to place students. This occurred frequently in Grade 1, where previous test data were usually lacking. Therefore, recommendations from kindergarten teachers, any readiness data and eligibility based on low-income status were used to place first-grade students. In general, supplemental means to identify students for experimental and control groups were used extensively in some districts with high mobility rates; however, the pre-drawn list of students formed the basis for selection decisions in most school districts.

Experimental Procedures

Evaluation of the effectiveness of the experimental programs has been accomplished through standardized testing conducted at the beginning and end of the 1970-71 school year. One standardized test was administered to experimental, control and other groups in each grade.* Pretesting took place within 10 days of the start of school and posttesting commenced 15 days before the end of school. For posttesting, an alternate form of the same pretest instrument was used.**

In addition, "follow-up" testing was accomplished in selected sites and grades within sites during the late fall of the current (1971-72) school year, using the same test instrument. Both previous experimental and control group students were tested in these sites. Sites and grade levels within each site where follow-up testing was accomplished, along with the criteria used for selection of these sites and grades, are given in the Results Section of the report.

Evaluation of the experimental programs is made mainly in terms of experimental versus control group test achievement. However, additional comparisons are made with the comparison and special treatment groups, when achievement data are available for these groups. Raw score units are used as the main basis for this evaluation as opposed to grade equivalence units. This is because raw score units, while less meaningful to the casual observer, do not possess psychometric distortions which might affect the results of statistical analyses.

---

* Comparison group posttesting was done only in selected sites and grades, due to the absorption of comparison group students into the experimental group, at several grade/sites. Sites and grades where comparison groups were posttested are given subsequently in the Results Section of the report.

** An exception to this occurred in Grade 1, as described subsequently in the section on Achievement Measures.

For purposes of payment to technology companies, experimental students only were administered the reading and math portions of an additional standardized test, on a pre- and post-basis. Each child was assigned randomly to take one of three tests in use for each grade and subject and took an alternate form of the same instrument for a posttest. Payment to the contractors was based upon experimental student gains shown between these pre- and posttests, measured in grade equivalents.

It is important to keep the distinction between evaluation and payment testing in mind. Evaluation testing is used to compare the achievement of experimental to control students, measured in raw score points on one set of instruments. The evaluation analysis is less in terms of gain for the experimental students than it is the comparison of achievement level for E and C students at the end of the year, with the pre-test being only one contributor to that status. Payment testing, on the other hand, measured gain on a pre-post basis in terms of grade equivalents for experimental students only as a basis for contract settlement. Battelle has not analyzed payment test results, since these results do not contribute to the evaluation analysis.

In addition to contractor payment based on gains in standardized tests, up to 25 percent of contract payments was based upon students' performance on interim tests oriented to the curricular objectives of each company. These tests were given at five times during the school year, in each grade. Although useful for payment purposes, the interim tests were not used for evaluation, because of the multiplicity of tests and objectives involved.

Finally, for evaluation Battelle's Columbus Laboratories distributed a Parent Questionnaire in early 1971 to all parents of the students in all groups. This questionnaire contained items concerning the attitudes of parents toward education in general and toward special education programs, the educational future of their children, and information about the educational, occupational and financial status of the parents. The results of this questionnaire are summarized in subsequent sections of this report. In addition to the Parent Questionnaire, information on the race, sex, age, and school attendance (1970-71 and 1969-70) of experimental and control students which had been kept by local project personnel was made available to Battelle for the evaluation analysis.

The Incentives Only School Districts

In order to provide an assessment of performance incentive contracting without the use of a private educational technology company, OEO contracted with two school districts to sponsor "Incentives Only" projects. The districts were Mesa, Arizona, and Stockton, California. Each of the districts subcontracted with its local educational association to test the use of incentives in regular classrooms. Apart from the incentives, all other aspects of the curricula and classroom projects were largely unchanged.

The local educational associations signed incentive contracts which specified that the amounts to be earned were dependent upon gains achieved on standardized tests used for payment. The disposition of funds earned, whether to reward the students or the teachers themselves, was left to the local educational association. In both Mesa and Stockton, most of the initial spending was to provide incentives to students, not to reward teachers. However, incentive payments to teachers will be made.

The hypothesis underlying these two programs was that low-income, low-achieving students instructed for one year by regular school techniques to which the use of incentives had been added would register better achievement on standardized tests than a similar control group of students not receiving the incentives. As with the eighteen technology company sites, the control students were usually designated in different school buildings in order to prevent any confounding of experimental versus control comparisons by having the effect of the experimental programs "rub off" or become assimilated through adjacent classrooms. The experiment was to be conducted within each of Grades 1, 2, 3, and 7, 8, and 9, with approximately 100 experimental and 100 control students in each grade. Experimental and control students within each grade were given standardized achievement tests at the beginning and end of the school year, in order to evaluate program impact on student achievement.

In the above basic matters of design, the two projects at Mesa and Stockton were similar to the projects in the other eighteen districts involved with private educational technology companies. Also, the standardized achievement measures used for evaluation and payment testing were the same as for the technology company sites, as well as the same basic testing design and procedures for administering these tests.*

The incentives only sites should not be considered comparable to the other eighteen in terms of techniques and results, however. The "Incentives Only" districts were selected later than the other eighteen districts and pretesting there took place more than one month after school began. Unlike the eighteen sites with performance contracts in which innovative hardware, software, and incentives were all being tested, the Incentives Only" districts were testing only the use of incentives added to regular class offerings. Unlike the eighteen sites where OEO and the schools contracted for the entire educational package, funds were provided only for the incentives, plus administrative expenses. For the above reasons, results and conclusions concerning experimental program impact at Stockton and Mesa are given in a separate section of this report. The reader is referred to another report** on Stockton and Mesa for a description of the experimental programs implemented at each site, description of the target population involved, a summary of pre- and posttesting conditions, and interim results and conclusions of statistical analyses.

---

* The one exception to this occurred at Mesa, where the junior high students did not receive a payment test, but only the evaluation test.

**Interim Report on the OEO Experiment in Educational Performance Contracting: The Incentive Only Sites. February 7, 1972. Battelle, Columbus Laboratories, 505 King Avenue, Columbus, Ohio, 43201.

DESCRIPTION OF TECHNOLOGY COMPANY PROGRAMS*

In this section, a general summary is given of the six technology company programs. To facilitate comparisons between the programs, each company program is described with respect to: (1) personnel and training, (2) curriculum and materials, and (3) use of incentives.

Alpha Learning Systems, Incorporated

Personnel and Training

Alpha trained and supervised teachers that were employed by the public schools, along with paraprofessionals from the community, to carry out their program. The company worked with teachers already assigned to the target schools selected for the experiment, training them in instructional and contingency management, flow charting, and the development of individualized programs.

Alpha was the only company utilizing teachers employed by the schools; in the other five companies, the teachers were employees of the company.

Curriculum and Materials

An oral-phonics approach was used for beginning reading. The program combined the Miami Linguistic Series, Language Master Series, McGraw-Hill Sullivan Series and some special materials developed by the Southwestern Educational Cooperative Laboratories. Once basic reading proficiency was established, greater emphasis was placed on written programs--e.g., SRA language labs and the Economy Series. The beginning math program consisted primarily of TMI-Grollier, progressing to Addison-Wesley. The junior high programs contained the same basic materials, but were expanded to include SRA's Reading for Understanding and Reader's Digest Skill Builders. Alpha used a wide range of individualized, self-paced materials, all of which are available commercially. Hardware was kept to a minimum.

---

* Summarized from the "Final Report to the Office of Economic Opportunity: Performance Incentive Remedial Education Experiment", August 31, 1971, Education Turnkey Systems, Inc., Washington, D.C. 20036.

Use of Incentives

An explicit system of contracting between the student and teacher for curriculum and available time was developed. Given a prepared range of tasks for one day, the student determined the order in which he performed the tasks and the reinforcements he would experience upon completion of a prescribed proficiency in each task. Alpha employed both intrinsic and tangible incentives with the students involved in the program. The former consisted of free time in which the student could do as he desired. A "free room" stocked with items the student enjoyed was made available. Greater levels of student learning efforts were rewarded by tangible incentives via macro contracts. These were in terms of money or additional free time.

Learning Foundations, Incorporated

Personnel and Training

Learning Foundations hired local individuals to serve as project "administrators" and program "supervisors". Paraprofessionals were used extensively to produce a supervisor-student ratio of one to five in the elementary school and one to seven at the secondary level. Originally, the "administrator", a certified teacher, was responsible for 600 students. Other certified teachers were added later in one site because of political and legal considerations.

Curriculum and Materials

Learning Foundations relied heavily on the use of teaching machines. The auto-tutor, variable-speed reader, EFI Card Reader, linear tutor, and self-pacing audio-visual projector comprised the hardware. The Craig program was used extensively for instructional and testing purposes. The Sullivan programs in reading and math were also used considerably in the Learning Foundations program. The other materials were primarily learner paced and geared to individualization of instruction. All the hardware is commercially available.

Use of Incentives

There was a heavy emphasis on incentive in the Learning Foundations program. Students were awarded points for achievement in areas of attendance, speed, and comprehension. These points were in the form of play money which could be used to "purchase" listed gift items. Some of the rewards further reinforced the learning process--e.g., dictionaries, telescopes, etc. In addition, bonus points were given for greater learning and attendance effort.

Plan Education Centers, Incorporated

Personnel and Training

Plan hired equal numbers of professionals and paraprofessionals locally for their learning centers. The Project Administrator was an individual hired specifically for this project. Preschool workshops and inservice workshops were held to train the staff on site.

Curriculum and Materials

Plan's learning centers were stocked with a variety of materials, the largest of any of the systems proposed. Use of hardware was minimal. In the Plan program, mathematics was taught as a language, as a statementgenerating system. By combining the teaching of reading and mathematics, especially at the early stages, and by interrelating vocabulary, sentence structure and comprehension, Plan coordinated learning in both areas. The Sullivan reading and mathematics materials along with SRA materials constituted the main components of the total program. The materials were largely learner paced and geared to individualization of instruction. There was some small group instruction, however, as Plan's program called for grouping by disability for intensive instruction for short periods. This was primarily tutoring by the paraprofessionals.

Use of Incentives

Plan did not provide teachers and students "tangible" rewards. Rather, through careful diagnosis and prescription, the student would be rewarded intrinsically as he progressed through the material. The extent to which rewards for students were provided by teachers did not differ significantly from those provided by other teachers in the system (e.g., verbal reinforcement, smiles, etc.)

Quality Educational Development, Incorporated

Personnel and Training

Local personnel were used as teachers and aides in the Q.E.D. program wherever possible. Individuals, not originally employed by Q.E.D., were hired specifically for this project as Project Administrators. An equal number of paraprofessionals and professionals were hired and used.

Curriculum and Materials

Two programs were used to teach beginning reading--at two sites, the Furguson ITA Program was used initially although later replaced; at the other, the Evans Reading Program. The Sullivan Reading Program was next in the progression. At the junior high school level, the Job Corps

Grading Reading Selections were added. In mathematics for both elementary and secondary levels, Houghton-Mifflin's Modern School Mathematics: Structure and Use Program was used in conjunction with the Stockton Individualized Learning System.

Q.E.D. based their instructional systems on programs of proven effectiveness although not widely used. This approach made optimum use of inexpensive hardware for the presentation of materials and provided a high teacher-machine to student ratio. A program of individually prescribed instruction was used consisting of a placement system, pre- and post-tests, and individualized materials enabling the child to progress at his own rate. As gaps appeared in the existing materials system, Q.E.D. created individualized, self-paced materials designed to modify or supplement the curriculum.

Use of Incentives

Contingency management as a direct application of reinforcement technology was utilized by Q.E.D. in the following way. The students were issued "credit cards" enabling them to receive points to their credit as they achieved various modes of success. They could then acquire articles from a "catalogue store" or "buy" time to pursue a preferred activity. Rolls of tickets as one might find at an amusement park were also used for this purpose.

Singer/Graflex

Personnel and Training

Local professionals and paraprofessionals were employed at a ratio of approximately five to two. Leadership and supervisory positions were filled by Singer/Graflex veterans with the majority of the project staff hired locally. An intensive training program for all staff at the beginning of the project was considered essential, and periodic training sessions were conducted throughout the term of the contract.

Curriculum and Materials

Singer relied heavily on the Job Corps Reading and Math Programs. The Sullivan Reading Program, SRA materials, and the Sullivan Math Program received special emphasis. In addition, the Singer Right to Read Program was used. Some audio-visual materials were utilized. These included filmstrips, tapes, controlled readers and language masters. Singer used commercially available materials in whole or in part depending upon the needs of the student. They were largely learner paced, based on individualized instruction.

Use of Incentives

The Singer approach called for intensive use of incentives. Achievement certificates and emblems were awarded. Students were encouraged to accumulate achievement points and bonus achievement points to exchange for catalogue merchandise. At the secondary level, achievement was rewarded with free time in which the student engaged in an activity of his choice.

Initially, social behavior was stressed and rewarded, e.g., proper use of the materials, cooperation with the instructors, etc. The emphasis then shifted to the reinforcement of appropriate learning behavior--the satisfactory completion of an assigned lesson. In the beginning, tangible items for all the students were used extensively. Later in the project, especially at the secondary level, intangible rewards such as free time to play games, movies, a trip for free hamburgers and milkshakes were stressed.

Westinghouse Learning Centers

Personnel and Training

The program staff was composed of a combination of certified teachers and paraprofessional aides. In addition, a Center Manager was employed at each location to serve as the overall supervisor. The center staff and aides were recruited from the local community. The professionals in some cases, were former school teachers, and the aides were chosen on the basis of their potential for establishing rapport with the student population. All staff members were trained in the operation of the instructional system by Westinghouse personnel. Periodic training was also conducted throughout the duration of the project.

Curriculum and Materials

To teach beginning reading, Westinghouse relied primarily on the BRL Reading Readiness Program. The Sullivan Reading Program followed and was used in conjunction with SRA labs and Charles Merrill materials as the main components of the reading program. For arithmetic, the Sullivan Basal Mathematics Program was heavily relied upon, supplemented by workbooks and audio-visual instructional materials when needed.

Westinghouse's instructional sequence intended to teach specific skills, using portions of commercially available materials. The curriculum included materials that were designed to be self-instructional, learner-paced, and those that taught the skills outlined in the curriculum. Also, the use of instructional hardware was observed in the Westinghouse system. For the most part, standard cassette tape players with headsets and reel-to-reel tape players were required for use with certain materials.

Use of Incentives

The contingency management motivational method employed in the operation of the Westinghouse program was characterized by contingency contracts--a daily agreement between the staff and the student stating precisely how many prescribed instructional sequences the student would complete in order to earn some free time. Implementation of this design involved the designation of two areas in the classroom--one area being a study area; the other, an activities area.

Tabular Summary

Some aspects of the experimental programs are summarized in tabular form in Table 1 on the following page. The table shows, for each technology company, percent of paraprofessionals employed, student-staff ratios, and type of incentives and instruction used. Both teachers and paraprofessionals are included in the student-staff ratio calculations.

TABLE 1. COMPARISON OF PARTICULAR ASPECTS OF EXPERIMENTAL PROGRAMS

Company	Average Percent of Paraprofessionals	Average Student/ Staff Ratio	Use of Incentives	Instruction
ALPHA	Elementary: 45% Junior High: 35%	Elementary: 1/14 Junior High: 1/15	Heavy emphasis on tokens which could be traded for a variety of prizes or free time at any time	Heavy use of programmed materials and individual instruction
Learning Foundations	Elementary: 100% Junior High: 100%	Elementary: 1/5 Junior High: 1/6	Heavy use of incentives utilizing a point system which students could use to purchase desired items	Used programmed materials and individual or group instruction
Plan Education Centers	Elementary: 50% Junior High: 50%	Elementary: 1/6 Junior High: 1/5	No "tangible" incentives used; teachers provided conventional incentives (e.g., verbal reinforcement)	Used primarily teacher instruction with some tapes and cassettes
QED	Elementary: 50% Junior High: 50%	Elementary: 1/13 Junior High: 1/13	Heavy reliance on all types of incentives utilization of a point system based on student performance	Heavy use of programmed materials with tape and cassette teaching machines
Singer/ Graflex	Elementary: 55% (Does not include McComb: 9%) Junior High: 32% (Does not include McComb: 6%)	Elementary: 1/20 Junior High: 1/20	Intensive use of incentives progressing from tangible rewards to intangible rewards	Primary reliance on programmed materials and group instruction
Westinghouse Learning Centers	Elementary: 80% Junior High: 80%	Elementary: 1/12 Junior High: 1/12	Heavy use of incentives in elementary grades, but few incentives used in junior high. Incentives awarded infrequently at random	Heavy reliance on programmed materials with tape and cassette teaching machines in individual and group instruction

DESCRIPTION OF THE TARGET POPULATION

The major purpose of this section of the report is to present descriptive information concerning some of the more important variables that characterize the socioeconomic status, social and home background, and achievement levels of students who participated in the experiment. Of particular concern are family income and achievement levels of students, since the experiment was directed toward "low achieving" students from "low income" families. However, descriptive data are also presented on the racial composition of students in the study, along with the educational and occupational status of the students' families, in order to afford the reader a more complete description of the type of student for which the results and conclusions of this study apply.

In this section of the report, achievement level data are taken from the pretest administration of the evaluation test, to characterize the entry level achievement of students participating in the study. The racial composition of students was obtained from student locator cards, a system of student characteristics maintained by the Project Directors at each site. Data on family income, and educational and occupational status of students' parents, were secured from a family background survey, consisting of a questionnaire sent to students' parents (described below).

Data are presented separately for experimental and control groups, so as to show differences between the two groups. The implecations of these group differences for the analysis, along with the methodology for taking into account group differences, is treated subsequently in the analysis sections of the report.

Data are presented for students who were in the experimental or control group for the entire school year*. These full-year students constitute the basis for the analysis. Sufficient achievement test data were not available for separate analyses of "dropouts", or students entering the experimental group program after its onset. However, descriptive data characterizing the initial entry achievement level of dropouts is given at the end of this section, to serve as a basis for comparing students who remained in the program the entire year with those who dropped out.

---

* Table A-1, Appendix A, gives the number of full year students, by site, group, and grade.

Securing Family Background Information

A family background survey was implemented during the month of December, 1970, for purposes of gathering parent attitudinal and socioeconomic data. A total of 27,018 participant identified questionnaires were delivered to each of the twenty project directors for program participants identified in Battelle's master file. An additional 100 questionnaires were sent to each site for student replacements to the OEO experiment who were not identified previously for Battelle in the original master list. Additionally, a Spanish version of the questionnaire was prepared by Battelle and sent to those sites requesting dual forms: Grand Rapids, Stockton, and Bronx.

The method of delivering questionnaires to parents was as follows. First, site Project Directors distributed the questionnaires to teachers of the students in the study. The students were then instructed to take the questionnaires home to their parents. Parents then filled out the questionnaire in accordance with enclosed instructions, and then returned the questionnaire directly to Battelle. As part of the instructions, parents were assured that their answers to the questions would be treated in a confidential manner, with only concerned Battelle staff members seeing their responses.

Two methods of follow-up were used. First, four days after the questionnaires were distributed to students, each student received a reminder card, which they were instructed to take home to their parents. The card "reminded" parents of the significance of the study, and again asked for their cooperation in responding to and returning the questionnaire to Battelle.

The second method of follow-up consisted of sending out a second wave of questionnaires to identified non-respondents, again using the "student delivery" system. This second wave was sent out approximately three weeks after the first questionnaire was to have been returned to Battelle.

Table B-1 (Appendix B) presents data relative to the return rate of parents, including all follow-up efforts. Data are presented for each site, and for each group and grade within each site. Return rates are calculated on the basis of the student population in the experiment at the time questionnaires were received at the school district.

Inspection of these data show variability of return rates between sites, grades, and experimental versus control groups. However, return rates on the order of 40-50 percent are not atypical, and represent a rate usually considered normal for the method of survey implementation used in this study. Nevertheless, this leaves missing data for at least 1/2 of the students' families. Return rates are discussed below, where rates are given for full-year students on the particular variables of family income, education, and occupation.

Appendix C presents a copy of the Parent Questionnaire sent to students' families. Items 32, 22, and 28 in the questionnaire were the basis for the income, educational, and occupational data.

Characteristics of Participants

Race, Education, Income, and Occupation

As stated previously in the "Description of Experiment" section, many different types of school districts were selected to be in the experiment, geographically scattered across the United States. This diversity is reflected in variations of student characteristics from site to site, expecially with regard to race of the student, and family income. For example, for race, there may be one, two, or three predominant races represented by students at a particular site, and the race(s) which predominate vary from site to site. Consequently, for race and income, descriptive information is given on a site by site basis only.

Educational and occupational characteristics of students' families also varied from site to site, but to a lesser degree than for income and race. Consequently, both site by site as well as data aggregated across sites are presented for these variables.

Race of Students. Table D-1 in Appendix D shows the percentage distribution of race, for each technology company site, and for experimental and control groups within a site. For each site, data given are for Grades 1, 2, 3, 7, 8, and 9 combined, for full-year students. Also given is the sample size on which these race data are based, i.e., the number of full-year students for which data are available. Finally, a "response rate" is given, which is the percentage of full-year students for which race data were available.

Response rates for the race data are generally high, most frequently being in the high 90's, as inspection of Table D-1 shows. However, for the control groups at four sites (Hartford, Philadelphia, Rockland, and Taft), race data were not available from the sites.

Inspection of these data show differences between experimental and control students in racial composition within sites. Thus, of the 14 sites for which race data were available for both experimental and control groups, for 10 of these sites the percentage of Whites in the control group is higher than the percentage of Whites in the experimental group. The differences in these percentages is often quite marked. For example, at Anchorage, 91 percent of the control group is White, whereas only 54 percent of the experimental group is White. As another example, at Hammond, the percentage of Whites in the control and experimental groups is 87 and 57 percent respectively.

Inspection of the data in Table D-1 also reveal appreciable variability in racial composition of the student population from site to site. This is illustrated in Table 2, which lists the predominant races for each site, for the experimental and control groups. "Predominant" races are defined by listing those races (or that race) that accounts for at least 85 percent of the students, by taking the most frequently occurring race, the next most frequently occurring race, etc., until at least 85 percent of the students are accounted for. Within the experimental and control groups, races are listed (from left to right) in the order of decreasing dominance.

Inspection of Table 2 shows that in eight sites, there was only one predominant race. In four of these eight sites (Dallas, Hartford, Jacksonville, and Philadelphia), the Black race predominated; in three other sites (Portland, Rockland, and Selmer), the White race predominated; in one site (Taft), a predominance of Mexican-Americans is shown.

In seven of the sites, Blacks and Whites predominated. These seven sites were Athens, Grand Rapids, Hammond, Las Vegas, McComb, Seattle, and Wichita. In Anchorage, Whites, Eskimos, and Blacks accounted for at least 85 percent of the students in the study; at Fresno, the population consisted mainly of Mexican-Americans and Whites; and at Bronx, the predominant races were Puerto Ricans and Blacks.

In summary, often marked differences appeared between experimental and control groups within a site in racial composition. Further, across sites, a diversity of racial composition of both minority and majority group membership was represented in this experiment, reflecting the diversity of school districts selected to be in the study.

Family Income. Table 3 shows the median total family income for parents of the experimental and control students, for each of the eighteen sites in Grades 1, 2, 3, 7, 8, and 9, combined. Also shown is the percentage of students' families having an income less than the national average, by site and group (E or C).

TABLE 2. PREDOMINANT RACES OF FULL-YEAR STUDENTS, BY SITE AND GROUP

Site	Experimental Group	Control Group
Anchorage	White, Eskimo, Black	White
Athens	Black, White	White, Black
Bronx	Pureto Rican, Black	Puerto Rican, Black
Dallas	Black	Black
Fresno	Mex. Amer., White	Mex. Amer., White
Grand Rapids	White, Black	White, Black
Hammond	White, Black	White
Hartford	Black	--
Jacksonville	Black	Black
Las Vegas	Black, White	White, Black
McComb	Black	Black, White
Philadelphia	Black	--
Portland	White	White
Rockland	White	--
Seattle	White, Black	White
Selmer	White	White
Taft	Mex. Amer.	--
Wichita	Black, White	White, Black

TABLE 3. MEDIAN TOTAL FAMILY INCOME AND PERCENTAGE OF FAMILIES HAVING AN INCOME LESS THAN THE NATIONAL AVERAGE FOR FULL-YEAR STUDENTS, BY SITE AND GROUP

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids	
	E	C	E	C	E	C	E	C	E	C	E	C
Median Income	10,565	14,896	4,666	5,187	6,111	5,454	3,034	2,475	6,263	6,821	6,692	7,852
% Below Nat'l Average (9,867)	41	17	84	79	83	88	95	97	83	78	80	76
Sample Size:	147	229	276	294	87	75	233	176	225	266	239	214
Response Rate:	46	62	57	60	19	13	45	34	43	50	45	44

	Hammond		Hartford		Jacksonville		Las Vegas		McComb		Philadelphia	
	E	C	E	C	E	C	E	C	E	C	E	C*
Median Income	8,833	8,903	4,222	5,400	4,281	3,771	9,312	8,000	2,636	4,272	3,842	4,666
% Below Nat'l Average (9,867)	61	61	90	87	90	92	56	52	97	84	87	91
Sample Size:	231	308	94	130	386	327	167	56	181	129	96	35
Response Rate:	44	47	19	24	70	61	37	19	42	41	19	8

	Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C
Median Income	6,333	8,170	6,717	--	8,461	9,409	6,219	5,062	2,806	2,882	7,350	7,941
% Below Nat'l Average (9,867)	74	68	74	--	58	54	81	89	97	98	68	62
Sample Size:	226	316	305	--	214	276	317	150	118	88	289	226
Response Rate:	44	54	51	--	43	50	55	54	32	18	43	49

* Data for Grades 7, 8, and 9 only.

Median incomes shown can be compared to the overall national median income of $9,867*, as of March, 1971. All median incomes are less than this nationwide median, with the exception of Anchorage, which shows a median of $10,565 for the experimental group, and $14,896 for the controls.**

However, median incomes vary widely from site to site, with lows occurring at McComb, Dallas, and Taft (Median incomes are on the order of $2,000 to $3,000 at these three sites). Other sites show median incomes varying between $4,000 and $8,000, generally.

The percentage of families having an income less than the national average also varies from site to site, although these percentages are generally high. Thus, of the 36 site/group combinations, 24 of the combinations show at least 75 percent of the families having an income less than the national average.

All the above cited descriptive data needs to be interpreted with some caution, in view of the response rates shown in the table, Response rates from Bronx, Hartford, Las Vegas, Philadelphia, and Taft are particularly low. For example, at Taft, in the control group, income data were available on only 18 percent of the full-year students.

The data in Table 3 reveal E/C group differences in income. For example, in Anchorage, the median income for the control group is on the order of $4,000 higher than for the experimental group. However, the magnitude and direction of group differences in income within a site is best illustrated in Appendix E, which gives median incomes by grade level, as well as by site and group. For example, at Jacksonville in Grades 1, 2, and 3, the experimental group has appreciably higher median incomes than the control group, whereas the reverse is true for the secondary level (i.e., the control group exhibits higher median incomes than the experimental group in Grades 7, 8, and 9). Differences such as these tend to be obscured in Table 3, both in terms of magnitude, and direction of group differences.

---

* Taken from "Current Population Reports: Consumer Income", U.S. Department of Commerce, Bureau of the Census. Series P-60, No. 78, May 20, 1971.

** Higher median incomes are to be expected in Anchorage, since the cost of living there is substantially higher than in the 48 States.

From an analysis point of view, the data given in Appendix E by grade are more descriptive, since, as described subsequently, the major unit of analysis is at a grade/site level. As another example, at McComb, families of elementary students in the control group have appreciably higher incomes than families of experimental group students, with much smaller differences between groups at the secondary level.

Father's Education. The percentage distribution of father's education for each site is shown in Appendix F, in three categories: completed less than high school; completed high school; completed more than high school. The latter category is an aggregate one, and includes some or completion of college, and/or some or completion of vocational school after high school, and/or some or completion of business school after high school.

Examination of these data shown certain E/C group differences within sites. Thus, in 9 of the 18 sites, the control group has a smaller percentage of fathers with "less than high school" than does the experimental group, and a higher percentage of "more than high school" than does the experimental group. For example, at Athens, 51 percent of the control group students have fathers with less than a high school education, whereas 64 percent of the experimental group falls in this category; 22 percent of the control group students at Athens have fathers with more than a high school education, compared with 10 percent in the experimental group.

A salient feature of the education data in Appendix F is the relatively large percentage of fathers having less than a high school education in both experimental and control groups. In general, this percentage is approximately 50, in each site. This is reflected in Table 4 on the following page, which shows that 50 percent of the fathers of students in the experimental group and 47 percent in the control group have less than a high school education, for all sites combined. Relatively few fathers have more than a high school education (22 percent in the experimental group, and 24 percent in the control group).

TABLE 4. PERCENTAGE DISTRIBUTION OF FATHER'S EDUCATION FOR FULL-YEAR STUDENTS, BY GROUP, FOR ALL SITES

Educational Level	Experimental Group	Control Group
• Less than High School	50	47
• Completed High School	28	29
• More than High School	22	24
	100	100
Sample Size:	3917	3332
Response Rate:	44	39

TABLE 5. PERCENTAGE DISTRIBUTION OF FATHER'S OCCUPATION FOR FULL-YEAR STUDENTS, BY GROUP, FOR ALL SITES

Occupational Category	Experimental Group	Control Group
• Semi or Unskilled	47	43
• Skilled Manual	28	30
• White Collar Professional or Business	25	27
	100	100
Sample Size:	2711	2428
Response Rate:	31	28

Father's Occupation. The percentage distribution of father's occupation is shown in Appendix G, for each site and group. Three categories are given: (1) semi or unskilled, (2) skilled manual, and (3) white collar, professional, or business. Semi or unskilled includes occupations such as cafeteria or laundry work, unskilled factory work, farmer's helper, factory machine operator, gas station worker, and foundry worker. Skilled manual work includes the skilled trades, such as welder, electrician, and applicance repairman. White collar, business or professional includes clerical or sales workers, small and large business owners or managers, and high and low paid professional work.

As summarized in Table 5, the preponderance of students' fathers across sites have occupations falling in the semi or unskilled category. Percentages are 47 and 43 percent, for the experimental and control groups, respectively. Approximately 30 percent of the fathers for the E and C groups fall in the skilled manual category, and approximately 25 percent fall in the white collar, business, or professional category.

As with the variables of education, income, and race, E/C group differences within sites occur. For example, at Selmer, the occupational "level" is higher in the experimental group than the control group. Thus 60 percent of the control student's fathers fall in the "semi or unskilled" category, whereas only 45 percent of the experimental student's fathers have occupations in this category, with a greater percentage of fathers in the white collar, professional, or business category.

The response rates for occupational data are somewhat lower than for income and educational data, due to the absence of some fathers from the home, and because some fathers were not currently employed. Thus, the descriptive data in Appendix G need to be qualified accordingly.

Achievement Levels

Tables H-1, H-2, H-3, and H-4 in Appendix H present data relative to the entry level of students in the experiment.

Tables H-1 and H-2 in Appendix H give mean grade equivalency (GEQ) values for reading (Table H-1) and for math (Table H-2), by site, group, and grade. For Grades 2, 3, 7, 8, and 9, the tables give the grade equivalency corresponding to the mean raw score on the evaluation pretest, for full-year students. For Grade 1, the tables given the stanine value* corresponding to the mean raw score on the evaluation pretest, again for full-year students.

Inspection of Tables H-1 and H-2 show that "average" grade equivalency values are almost universally below grade level, across sites and E and C groups, with the amount below grade level (the grade level "decrement") generally increasing with grade level. However, although students are generally below grade level in both experimental and control groups, differences between E and C groups in entry level are nonetheless apparent, as examination of Tables H-1 and H-2 show. In this connection, at the secondary level, average grade equivalencies are generally higher for the control groups. Thus, of the 106** site/grade/subject area combinations at the secondary level, the control group has a higher grade equivalency entry level in 84, or approximately 80 percent of these combinations. At the elementary level, although group differences often exist, the tendency for the control group to have higher entry levels than the experimental group is much less marked (the control group has a higher entry level in 66 of the 106 possible grade/site/subject area combinations at the elementary level).

Tables H-3 and H-4 in Appendix H give percentage of students below grade level, for Grades 2, 3, 7, 8, and 9, for each of reading and math. For example, for Anchorage, E group, Grade 8 in reading, 91 percent of the students registered below grade level on the evaluation pretest. Inspection of these tables show very large percentages of students below grade level. There are, however, some very few exceptions to this in certain sites and grade/groups within sites, as inspection of Tables H-3 and H-4 show.

---

* A stanine is a transformation of a raw score, ranging from 1 to 9, such that the lowest and highest 4 percent of the raw scores are assigned the values 1 and 9, respectively; the next lowest and highest 7 percent are 2 and 8; the next lowest and highest 12 percent are 3 and 7; the next lowest and highest 17 percent are 4 and 6; and the middle 20 percent are 5.

** (18 sites) (3 grades) (2 subject areas) = 108 grade/site/subject combinations at the secondary level, less 2 combinations in Grade 9 in Bronx, where data were not available.

The very large percentages of students below grade level are summarized in Table 6 below. This table shows percentage of full-year students below grade level by grade, subject area, and group, for all 18 sites combined. Percentages below grade level are generally in the mid 80's to high 90's. The smallest percentages are 83 (Grade 9, reading; Grade 2, mathematics) and 84 (Grade 2, reading).

TABLE 6. PERCENTAGE OF FULL-YEAR STUDENTS BELOW GRADE LEVEL FOR READING AND MATHEMATICS, BY GROUP AND GRADE, FOR ALL SITES

	Reading		Mathematics	
Grade	Experimental	Control	Experimental	Control
2	91	84	85	83
3	97	86	95	86
7	96	88	96	91
8	95	91	96	92
9	93	83	97	90

The entry level achievement of students in the experiment is also summarized in Table 7, in grade equivalence units, for all eighteen sites combined. For Grades 2, 3, 7, 8, and 9, the table gives the grade equivalence corresponding to the mean raw score for the eighteen sites on the evaluation pretest, for reading and mathematics and for experimental and control groups. Stanine values are given for Grade 1.

The data in Table 7 are portrayed graphically in Figure 1, where mean pretest grade equivalencies (GEQ's) are plotted for each grade level, group, and subject area (reading or mathematics). Grade 1 is not plotted, since data are not available in grade equivalence units for the Grade 1 evaluation pretest.

As shown in Figure 1, both experimental and control students in Grade 2 start out at about the same amount below grade level (as indicated by the vertical distance from the 45 degree line), and both

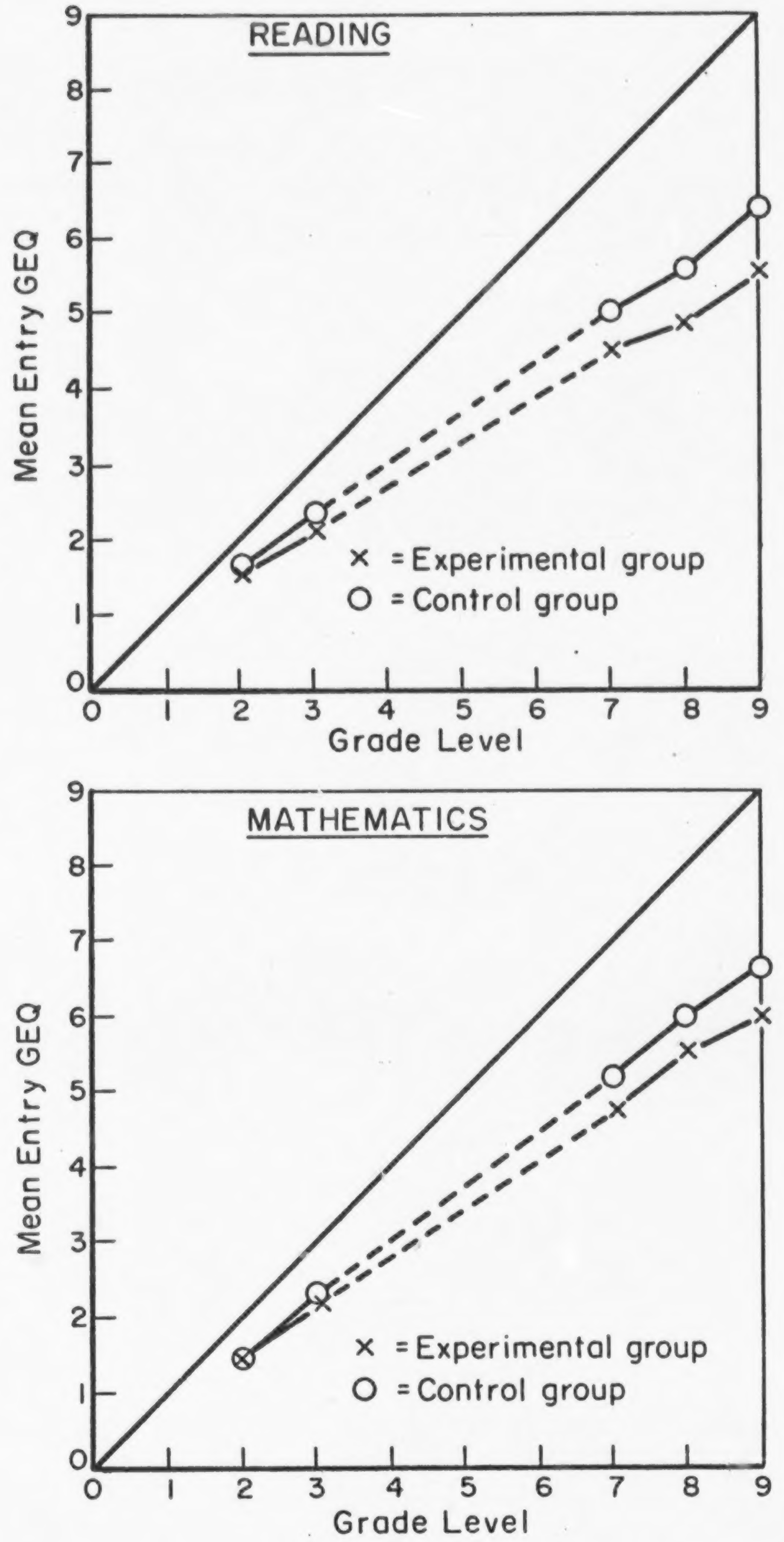


FIGURE 1. OVERALL AVERAGE PRETEST GRADE EQUIVALENTS FOR EXPERIMENTAL AND CONTROL GROUPS, FOR EACH GRADE IN READING AND MATHEMATICS.

E and C groups get further behind grade level with each succeeding year in school. Also, as the grade level increases, the separation between experimental and control groups increases, and more so for reading than for mathematics. For each group in each subject area, the rate at which students fall behind grade level seems to be about constant from year to year, as indicated by the fact that each of the four plots is essentially a straight line. Finally, both E and C students fall further behind grade level in reading than in mathematics, as shown by the lower "growth rates" (or smaller slopes of the lines) for E and C groups in reading than in mathematics.

TABLE 7. OVERALL MEAN PRETEST GRADE EQUIVALENT VALUES FOR READING AND MATHEMATICS, BY GROUP AND GRADE, FOR ALL SITES

	Reading		Mathematics	
	Experimental	Control	Experimental	Control
Grade 1	2*	3*	2*	3*
Grade 2	1.5	1.6	1.4	1.4
Grade 3	2.1	2.3	2.2	2.3
Grade 7	4.5	5.0	4.7	5.1
Grade 8	4.8	5.6	5.4	5.9
Grade 9	5.6	6.4	6.0	6.6

* Stanine values for Grade 1

Other Variables

In addition to the variables of race, income, education, occupation, and pretest, experimental and control groups were examined with respect to several other variables, many of which were attitudinal items from the Parent Questionnaire. These other variables were examined largely from the point of view of comparability of E and C groups, and thus the focus was on examining how the two groups differed on these variables.

Additional items from the Parent Questionnaire that were examined in order to study group differences are as follows:

- Importance of schooling (Item 2)
- Approval of new teaching method for own children (Item 10)
- Assistance in doing school work (Item 18)
- Encouragement to do well in school (Item 19)
- Amount of schooling desired for child (Item 20)
- Amount of schooling expected for child (Item 21)
- Employment status of husband and presence/absence of husband in household (Item 26)
- Per capita income (derived from Items 30 and 32)
- Welfare experience (Item 34).

The Parent Questionnaire, with the full statement of each of the above items, is presented in Appendix C.

In addition to the above items, data were obtained from student locator cards for student age, and for student attendance during the 1970-71 school year. Experimental versus control group differences were also examined with respect to these two variables.

Group differences on all of the above variables were studied in the following manner. For any given variable at a given grade/site, an examination was made of the variable's distribution for the E and C groups, for Grades 1, 2, 3, 7, 8, and 9 combined. If the E group's relative frequency distribution differed by more than 10 percentage points from at least one corresponding category in the C group's distribution, the groups were judged to differ on the variable. For example, at Dallas for Item 34 (welfare experience), 67 percent of the control group's families reported they were currently receiving financial aid, whereas only 46 percent of the experimental group reported that they were currently receiving aid. Thus, at Dallas, the E and C groups were judged to differ on this item.

The results of this analysis are shown in Table 8. An "X" appears wherever an E/C difference exists according to the criterion described in the previous paragraph.

The table also contains the variables of race, total family income, father's education, and father's occupation, as well as the variables mentioned immediately above, in order to provide a summary of group differences for all variables.

Of the attitudinal items analyzed from the Parent Questionnaire (Items 3, 10, 18, 19, 20, and 21), of particular interest is the number of sites on which E and C groups differed on Item 10, where parents were asked how much they approved (or disapproved) of their children being taught by a new teaching method. Parents responded in four categories ranging from "approve very much" to "disapprove very much".

As shown in Table 8, of the 17 sites where data were available, an E/C group difference on Item 10 was judged to exist in 13 of the 17. Further, in several sites, the families of the experimental group had a more favorable attitude toward their children being taught by a new method than did families of the control group. This is illustrated in Appendix I, which shows, for each of the 18 sites, the frequency distribution of responses to each of the four categories in Item 10. As can be seen there, in several sites, the category "approve very much" is responded to more frequently in the experimental group than the control group.

As mentioned previously, the method for taking into account group differences as described above in assessing program impact on posttest performance is presented subsequently in the analysis section of the report.

Comparison of Full-Year Students and Dropouts

In any study where experimental group members are subject to attrition, it is of interest to inquire whether subjects who "dropout" of the experimental treatment are systematically different from those who remain for the entire treatment. This is of particular concern in this experiment, since, as mentioned earlier, the analysis is confined to those students who remained in the program the entire year.

In this study, most students are not old enough to drop out of school. Thus, the reason for a student dropping out of the program can generally be attributed to other factors, such as the student's parents moving from the school district.

TABLE 8. ADDITIONAL VARIABLES, OTHER THAN PRETEST, ON WHICH THE EXPERIMENTAL AND CONTROL GROUPS WERE JUDGED TO DIFFER FOR EACH SITE

Variables	Sites SELMER	ATHENS	WICHITA	DALLAS	ANCHORAGE	ROCKLAND	LAS VEGAS	FRESNO	PHILADELPHIA	TAFT	GRAND RAPIDS	HARTFORD	MCCOMB	SEATTLE	PORTLAND	JACKSONVILLE	HAMMOND	BRONX
Importance of Schooling (2)*						ND												
Approve New Teaching Method (10)	X		X		X	ND	X	X		X		X	X	X	X	X	X	X
Assist School Work (18)						ND			X	X		X						
Encourage to Do Well (19)						ND				X								
Amount of School Desired (20)				X		ND						X		X				
Amount of School Expected (21)			X			ND							X	X	X			
Father's Education (22)		X		X	X	ND					X		X			X		
Father's Job Status (26)				X		ND			X	X	X		X					
Father's Occupation (28)	X	X			X	ND	X		X					X				
Total Family Income (32)					X	ND	X		X				X					X
Per Capita Income (Derived)	X			X	X	ND				X	X	X	X	X	X		X	X
Welfare Experience (34)				X	X	ND	X			X	X	X	X	X	X			X
Student Age					X	ND			ND	ND		ND		X	ND	X		X
Student Race		X	X		X	ND		X	ND	ND		ND	X	X			X	
Attendance (1970-1971)				X		X			ND	ND	ND		X		X			ND

*The number in parentheses is the actual item number in the Parent Questionnaire.

X = E/C difference greater than 10 percentage points in at least one of the categories of the variable.

ND = No Data

In order to compare full year experimental students with experimental students who dropped out of the program at some point during the school year, the pretest scores of both groups were examined. In particular, the mean pretest scores of full-year students was compared to the mean pretest scores of dropouts, aggregated across the 18 sites. This was done for each grade and for each of reading and mathematics within a grade.

Results of this comparison are shown in Table 9 below. Raw score means are given for full-year students and dropouts, along with associated grade equivalence values. Stanine values are given for Grade 1.

TABLE 9. MEAN PRETEST VALUES (AND ASSOCIATED GRADE EQUIVALENTS) FOR FULL-YEAR EXPERIMENTAL STUDENTS AND DROPOUTS, BY GRADE AND SUBJECT AREA

	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math
Full-Year Students	68 (2)*	68 (2)*	32 (1.5)	28 (1.4)	33 (2.1)	44 (2.2)	40 (4.5)	43 (4.7)	32 (4.8)	39 (5.4)	38 (5.6)	46 (6.0)
Dropouts	68 (2)*	68 (2)*	30 (1.4)	28 (1.4)	33 (2.1)	46 (2.2)	40 (4.5)	42 (4.6)	32 (4.8)	39 (5.4)	37 (5.5)	44 (5.9)
Number of Full-Year Students	1047	1042	1242	1152	1317	1242	1211	1197	1151	1065	1093	1102
Number of Dropouts	196	196	208	190	185	174	179	186	216	214	224	222

*Stanine values.

Inspection of Table 9 shows remarkably similar pretest means for full year students and dropouts. In terms of grade equivalencies, full-year students and dropouts never differ by more than one-tenth of a year, and the grade equivalencies (or stanines for Grade 1) are identical for the two groups for eight of the twelve grade-subject combinations. In terms of raw score means, the full-year students and dropouts never differ by more than two raw score units, and the mean raw score values are again identical for the two groups for eight of the twelve grade-subject combinations.

Thus, it is concluded that the overall entry level achievement of full-year students and dropouts is the same, and that the dropouts are not a different group than full-year students with respect to achievement level.

ACHIEVEMENT MEASURES*

Battelle, in its role as Test and Analysis Contractor, had responsibility for advising the Office of Economic Opportunity in the selection of standardized achievement tests for administration to students involved in the experimental program. In this section of the report, a summary is given of the criteria used for the selection of the evaluation tests, along with an identification and brief description of the evaluation tests selected. Identification and descriptions of the certification tests are given in Appendix J of this report.

Selection of Evaluation Tests

The criteria for the selection of the tests included the following:

- The norms for the tests should be based on a relatively recent national standardization sample having a reasonably large number of students from large metropolitan and rural school district poverty areas as well as those from the more "average", middle class school districts.
- A relatively current revision of test content that included subtests which sample a wide range of skills and knowledges of high relevance to the generally accepted goals and objectives of the curriculum.
- A high degree of reliability.
- Clear and simple directions for administering the test.
- The availability of alternative forms.

In addition to the above, it was decided that the tests should (1) to the extent possible, consist of a single battery of tests having several levels appropriate for students enrolled in the six grades and (2) reflect the generalized outcomes of the overall curriculum design for these grades.

---

*Abstracted from the "Summary Report on Description of Test Selection Rationale to Office of Economic Opportunity", January 12, 1971, Battelle's Columbus Laboratories, Columbus, Ohio 43201.

The Battelle staff met with a representative of the Office of Economic Opportunity to select from a listing of various achievement test batteries, those tests needed to meet the requirements of the program. The Office of Economic Opportunity representative had researched existing standardized tests throughout the summer months and presented his recommendations, as one input to the selection process. Each of the tests was reviewed, with respect to the criteria previously described. For this review, information was obtained from the following sources: (1) knowledge gained by the Battelle staff through direct use of the tests in other testing programs (2) technical manuals and test construction information (3) Buros' "Sixth Mental Measurements Yearbook", and (4) a report prepared by UCLA's Center for the Study of Evaluation, entitled "Elementary School Evaluation Kit".*

In order to assure that most students would achieve some degree of successful performance on a test, and to reduce the dimension of test difficulty for potential low-achieving students, it was decided, for a given grade, to use a level of a test that had an intended grade range that included the previous grade. For example, instead of seeking a test for ninth grade students that was designed to be given at Grades 9 through 12, a test was sought having an intended grade range of, say, 7 through 10, or 8 through 11.

Table 10 presents a listing of the evaluation tests selected for each grade, along with the intended grade ranges. Two tests were selected for the first grade. For Grades 2, 3, 7, 8, and 9, the Metropolitan Achievement Test Series (1970) was selected, with the test levels used for each grade as indicated in the table. For Grades 8 and 9, the same test level (Advanced) was used in each grade.

---

*ELEMENTARY SCHOOL EVALUATION KIT, prepared by the Staff of the Elementary School Evaluation Project, Center for Study of Evaluation, UCLA, Developmental Copyright, 1970.

TABLE 10. LIST OF TESTS FOR EACH GRADE

rade	Title of Tests	Intended Grades
1	Stanford Early Achievement-Pretest	Kindergarten to 1.5
	California Achievement Test-Posttest 1970 Edition (Level 1)	1.5 - 2.0
2	Metropolitan Achievement Test 1970 Edition (Primary I) (Forms F and G)*	1.5 - 2.4
3	Metropolitan Achievement Test 1970 Edition (Primary II) (Forms F and G)*	2.5 - 3.4
7	Metropolitan Achievement Test 1970 Edition (Intermediate) (Forms F and G)*	5.0 - 6.9
8	Metropolitan Achievement Test 1970 Edition (Advanced) (Forms F and G)*	7.0 - 9.5
9	Metropolitan Achievement Test 1970 Edition (Advanced) (Forms F and G)*	7.0 - 9.5

*For each of Grades 2, 3, 7, 8, and 9, Form F was used as the pretest, and alternate Form G as the posttest.

Descriptions of Evaluation Tests Selected

Grade 1

Because of the scarcity of standardized achievement tests having a good balance of both readiness and reading items, and having norms on the performances of kindergarten and entering first grade children, it was possible to consider only two tests for measuring first grade entry-level achievement; The California Achievement Test, Level I, 1970 Edition and the Stanford Early Achievement Test. Of these two, the Stanford Early School Achievement Test was selected for use in measuring the achievement of first grade students in the pretest phase of the experiment, primarily because of the large number of readiness items it contains for measuring achievement. Another consideration, however, had to do with the design of the normative study. The study included more than 8,000 students in kindergarten and more than 11,000 students in first grade. A good balance of students from cities of over 100,000 population; cities from 10,000-100,000 population; urban areas below 10,000 population; and rural areas was also obtained. Communities were selected on the basis of median family income and median years in school of persons age 25 and above. In each case the median for the norm group was very similar to that of the United States as reported in the 1960 census.

Table 11 at the end of this section presents the reliability estimates, and total number of test items for the Stanford test. The reliabilities reported are acceptable for this grade level in comparison to the number of items in each subtest. This test is particularly attractive to the beginning first grade student. It is administered in five fairly short sittings during which the students answer questions aimed directly at assessing experience with words, numbers, pictures, and sounds. These provide the basis for assessing the attainment of more specific instructional goals at a later time.

The California Achievement Test, Level I, was considered the most satisfactory test for measuring end-of-year achievement of first grade students. The booklet is attractive and the printed pages and other materials are quite clear. It has a good balance between readiness and achievement items related to vocabulary, comprehension, and number skills.

The normative information is sound and is based upon a stratification of school districts according to a geographic location, average enrollment per grade, and community type (determined by size and density of the community). Data are currently being analyzed, and at this time it appears that the groups are properly represented in the sample in proportion to their approximate ratio within the total population.

Kuder-Richardson reliability estimates for the California test (reported at mid-year) are shown in Table 12 for each subtest. These reliabilities reflect a problem with the reading comprehension sub-test at this level (r = .759). However, in the analysis, total reading scores are analyzed, consisting of both the comprehension and vocabulary subtests. Total reading, then, does have a satisfactory reliability (r = .95).

Grades 2, 3, 7, 8, 9

The 1970 Edition of the Metropolitan Achievement Test Battery was selected as the evaluation instrument for Grades 2, 3, 7, 8, and 9. This battery was selected primarily because it was one of the most recently revised in terms of content and normative data.

The format of each test is excellent and the subtests are relatively short compared to similar achievement batteries. There is no compromise, however, in the coverage of content, and at the higher levels, science and social studies are included in addition to the traditional coverage of reading, spelling, language, and mathematics skills. Subtests included in the battery for Grades 2, 3, 7, 8, and 9 are listed below.

Grade 2:

- Word Knowledge
- Word Analysis
- Reading
  - Part A: Sentences
  - Part B: Stories
- Mathematics
  - Part A: Concepts
  - Part B: Computation

Grade 3:

Word Knowledge
Word Analysis
Reading
  Part A: Sentences
  Part B: Stories
Spelling
Mathematics Computation
Mathematics Concepts
Mathematics Problem Solving

Grades 7, 8, and 9

Word Knowledge
Reading
Language
Spelling
Mathematics Computation
Mathematics Concepts
Mathematics Problem Solving
Science
Social Studies

The tests of main interest in the analysis, reading and mathematics, contain subtests which measure the skills as follows:

- Word Knowledge measures extent of students' reading vocabulary
- Word Analysis measures student's knowledge of sound-letter relationships or skill in decoding
- Reading Measures students' comprehension of written material
- Mathematics Computation measures students' ability to compute
- Mathematics Concepts measures students' understanding of important mathematical principles and relationships
- Mathematics Problem Solving measures students' ability to apply knowledge in solving numerical problems.

The battery offers a good measure of the generalized outcomes of education in the United States. The tests included in the battery are of high technical quality with the estimates of reliabilities shown below in Tables 13 through 16, and the normative study includes many of the largest school districts in the country.

Certification Tests

Certification tests were selected using the same criteria as for the evaluation tests. These criteria were met for the certification tests, with the exception of one of the Grade 2 and Grade 3 reading tests (The ETS Survey of Primary Reading Development used as one of the certification tests in Grades 2 and 3). In this instance, the particular criterion not met was that of having norms based on a national standardization sample of students from large metropolitan and rural school districts. However, the standardization sample, although geographically limited, was heterogeneous in its characteristics, and thus considered appropriate.* Moreover, the ETS tests did meet all other selection criteria, and they were highly rated in a report which evaluated existing nationally standardized achievement tests, prepared by UCLA's Center for the Study of Evaluation.**

A listing of all certification tests used is given in Appendix J, for each grade and subject area. A description of these tests is also given there, covering the same items as previously described for the evaluation tests.

---

*See Appendix J for a description of the norming sample used.

**Op. Cit., page 44.

TABLE 11. RELIABILITY COEFFICIENTS AND NUMBER OF ITEMS FOR THE FIRST GRADE PRETEST (STANFORD EARLY ACHIEVEMENT)

Subtests	Reliabilities	Number of Items
The Environment	.82 (Split-half)	42
Mathematics	.82 (Split-half)	28
Letters and Sounds	.89 (Split-half)	28
Aural Comprehension	.77 (Split-half)	28

TABLE 12. RELIABILITY COEFFICIENTS AND NUMBER OF ITEMS FOR THE FIRST GRADE POSTTEST (CALIFORNIA ACHIEVEMENT TEST)

Subtests	Reliability KR20	Number of Items
Reading Total	.950	116
Vocabulary	.953	--
Comprehension	.759	--
Mathematics Total	.956	87
Computation	.947	--
Concepts and Problems	.904	--

TABLE 13. RELIABILITY COEFFICIENTS AND NUMBER OF TEST ITEMS FOR THE SECOND GRADE TEST (MET 70, PRIMARY I)

Subtests	Reliabilities KR20	Number of Items
Word Knowledge	.88	35
Word Analysis	.90	40
Reading	.95	42
Total Reading	.96	117
Total Mathematics	.93	62

TABLE 14. RELIABILITY COEFFICIENTS AND NUMBER OF ITEMS FOR THE THIRD GRADE TEST (MET 70, PRIMARY II)

Subtests	Reliabilities KR20	Number of Items
Work Knowledge	.93	40
Word Analysis	.90	35
Reading	.93	44
Total Reading	.96	119
Mathematics Computation	.86	33
Mathematics Concepts	.85	40
Mathematics Problem Solving	.88	35
Total Mathematics	.95	108

TABLE 15 . RELIABILITY COEFFICIENTS AND NUMBER OF TEST ITEMS FOR THE SEVENTH GRADE (MET 70, INTERMEDIATE)

Subtests	Reliabilities KR20	Number of Items
Word Knowledge	.92	50
Reading	.93	45
Total Reading	.96	95
Mathematics Computation	.84	40
Mathematics Concepts	.88	40
Mathematics Problem Solving	.89	35
Total Mathematics	.95	115

TABLE 16 . RELIABILITY COEFFICIENTS AND NUMBER OF TEST ITEMS FOR THE EIGHTH AND NINTH GRADE TEST (MET 70, ADVANCED)

Subtests	Reliabilities KR20	Number of Items
Word Knowledge	.93	50
Reading	.92	45
Total Reading	.96	95
Mathematics Computation	.91	40
Mathematics Concepts	.90	40
Mathematics Problem Solving	.90	35
Total Mathematics	.96	115

TESTING PROCEDURES AND CONDITIONS

In this section of the report, a summary is given of testing procedures and testing conditions for both the pre- and post-testing phases. Topics covered include test administration design, selection and training of test coordinators, implementation of the testing, and test coordinator reports on testing conditions.

Pretesting

Administration Design

The test administration design centered on a program of testing which included two minimal conditions as follows.

- All primary grades were to have the tests administered in morning sessions, in class sizes of 35 or less in Grades 2 and 3, and a class size of 25 or less in Grade 1.
- All junior high school students were to have the tests administered in morning sessions with the exception of one separately timed test administered in the afternoon. Class sizes were to be 100 students or less, with each group having one proctor for every 50 students in addition to the test examiner.

The two administrations, evaluation and payment (or certification) testing, were organized consistent with these minimal conditions and a testing schedule was devised for each. The evaluation testing called for two consecutive morning sessions for the elementary grades and two consecutive morning sessions plus one test administered on two consecutive afternoons for the junior high school grades. Evaluation tests were always administered before the certification tests (for both pre- and post-testing). The rationale for having the evaluation tests administered first was to preclude the possibility of introducing "practice effects" as a source of bias in the overall evaluation. The evaluation tests were administered to the experimental, control, and comparison groups, and, at some sites, to special treatment groups of students.

The second test administration, certification or payment testing, involved the administration of different batteries of tests to only the experimental students in order to obtain test scores which would serve as a basis of payment for the technology contractors. The certification test batteries were used to measure reading and mathematics skills only. The certification pretesting called for two consecutive morning sessions for the elementary grades and one morning session plus one separately timed test in the afternoon for the junior high school grades. One of three different test batteries was randomly assigned and administered to one-third of the experimental students in each of Grades 2,3,7,8 and 9. Only one certification test battery was administered to the students in Grade 1.

The recommended testing schedules were designed to eliminate the effects of fatigue, limited span of attention, and satiation, especially in the administration of tests for the elementary grades. The class size recommendations previously cited were developed to provide for control of the test administration to assure standard conditions, to preclude cheating, and to speed the administration of the tests.

Based on the recommended class sizes, and the targeted number of students to be tested in the various grade/groups at each site, estimates were made of the numbers of test examiners and proctors required to administer the tests at each site, for both evaluation and certification testing. For the primary grades, plans were made to recruit these examiners and proctors from among certified substitute teachers at the site. For the junior high school level, it was recommended that guidance personnel at the site be used to administer the tests, since such personnel are most familiar and experienced in test administration. Teachers were to be used in the event that guidance counselors could not be recruited.

In addition, it was planned that the Battelle test coordinators would provide a 1-day training session for the test examiners in the procedures for administering the tests. This training would be given to assure standardization of test administration procedures. Special training materials were developed for and provided to the Battelle test coordinators to assure some standardization in the training of test examiners. The training materials, in addition to outlining all aspects of the program and activities the coordinators would undertake, included handouts for test examiners which outlined their responsibilities and the principles of standardized test administration.

Selection and Training of Test Coordinators

All of the test coordinators selected by Battelle-Columbus to coordinate the test administration at the sites had at least a Master's Degree in psychology or a field of education, and had completed a course in tests and measurement. Of the 17 test coordinators who visited 20 sites, 5 had Doctorates and 12 had Master's Degrees. Six of the test coordinators were full-time Battelle-Columbus staff members, and 11 were consultants of Battelle-Columbus. Seven of these consultants are engaged full time in testing programs or services involving testing. Thus, all of the test coordinators were professionally qualified to coordinate the pretest administration at the project sites.

The test coordinators for the pretesting phase were briefed during a 1-day session at Battelle's Columbus Laboratories. The training consisted of an overview of the organization of the experimental program, identification of the groups of students to be tested at each site, and the specific responsibilities of Battelle-Columbus. Finally, the test coordinators were instructed as to the prearrangements made prior to their arrival on site, responsibilities they would have on site, and specific information needed by them to complete the pretesting arrangements. Emphasis at the session was directed toward the establishment of a consistent pattern of operation among the various testing sites. Such an emphasis was necessary to provide as much standardization in the administration of the tests as possible.

Pretest Implementation

The Battelle test coordinator arrived on site 3 days prior to the administration of the evaluation tests to review with the Project Director the organizational plans for conducting the test administration, and to observe physical facilities. Also, during this time, test materials arrived at the site, and the coordinator inventoried materials received against materials required. Although in several cases, difficult logistical problems were encountered in getting testing materials to the sites, the materials arrived at each site in time and in sufficient quantities to test the targeted number of students, with few minor exceptions that were rectified immediately by sending additional materials.

Problems concerning the selection and training of test examiners at the sites were minimal. Also, test coordinators reported that the cooperation of school personnel directly involved in the study was excellent. A few instances were recorded in which some of the staff members at certain sites presented problems, but none of them seriously threatened the pretest administration and most were resolved prior to the administration of the tests.

There were, however, several problems encountered at the sites in organizing and implementing the testing program, especially in the areas concerning student selection, to a lesser extent arrangement of physical facilities, and difficulties in adhering to the testing schedule. These problem areas were outlined in a previous report to OEO*. Of concern to this report, however, are assessments of classroom testing conditions at the sites, since such assessments bear directly on interpretation of the results and conclusions subsequently presented in the analysis section of this report.

Reports of Testing Conditions

The Battelle test coordinators were asked to report any event that occurred during the actual administration of the tests, which, in their opinion, might endanger the validity of test results. In this reporting, they were asked to be specific about grade level and testing group whenever necessary. A summary of these reports is given below, by site, for the evaluation pre-testing.

---

*"Pretesting for the Office of Economic Opportunity Performance Incentive Experiment in Education", January 25, 1971, Battelle's Columbus Laboratories, Columbus, Ohio, 43201.

Grand Rapids and Bronx. At two of the 18 sites (Grand Rapids and Bronx) reported events and testing conditions during the pretesting were of such a nature that a decision was made by Battelle and OEO to retest certain grade /groups, in order to obtain more valid test data. At Grand Rapids, severe reservations were expressed by the site Project Director and the test coordinator concerning the validity of Grades 7, 8, and 9 evaluation testing due to discipline problems, testing facilities, and difficulty of students in hearing test administration directions in the facilities used. Consequently, it was decided to retest these students, and the second testing went smoothly.

At Bronx, the second testing effort was much more extensive, since a greater number of grades and testing groups experienced unusual circumstances during the original testing. Retesting occurred for all grade/groups except Grades 1, 2, and 3 control groups. Retesting was judged essential due to a combination of several factors, including inability to get students identified for testing, inadequate and crowded physical facilities, humid weather, and student fatigue and disruptive behavior.

Conditions at Bronx were greatly improved on the second testing. There was a problem, however, in that the number of students tested in the Grades 7, 8, and 9 control groups was less than the planned number, on the retesting.

Reports of evaluation pretesting conditions for the other 16 technology company sites are summarized in Table 17. As can be seen from the table, for Anchorage, McComb, Rockland, Selmer, and Taft, no incidents were reported, for any grade, group, or subject matter area.

TABLE 17. TESTING INCIDENTS REPORTED FOR PRE-EVALUATION TESTING

Site	Incidents or Conditions
Anchorage	• None reported
Athens	• Grades 7, 8, 9, most experimental groups, reading: discipline problems
Bronx	• None reported on retesting
Dallas	• Junior High: Organizational confusion in getting testing underway, and in selecting students, that might have affected students test taking behavior
Fresno	• Grades 7, 8, 9, experimental and control, reading and mathematics: discipline problems, students not paying attention, some testing groups larger than planned
Grand Rapids	• Grades 2 and 3, control, reading and mathematics, for approximately 20% of students: noise, cheating • No incidents reported on Grades 7, 8, 9 retesting
Hammond	• Junior High: Low motivation and difficulty in getting students interested
Hartford	• Elementary and Junior High: Discipline problems; not following directions well; tendency of students to put little effort into the test taking
Jacksonville	• Grade 7, experimental, reading and mathematics: poor handling of group by examiner; noisy outbursts, test marking without question reading

TABLE 17. (Continued)

Site	Incidents or Conditions
Las Vegas	• Grades 1, 2, 3, experimental and control, reading and mathematics (more than 50% of students): crowded conditions and testing of two or three grade levels in single large room; discipline problems and controlling students; noise and talking from students • Grades 8 and 9, experimental and control, reading and mathematics: continued talking; discipline problems
McComb	• None reported
Philadelphia	• Extreme heat (95 F) during testing; rooms not air conditioned
Portland	• Grades 7, 8, 9, experimental and control, reading (all students): difficulty in controlling students; over crowding in testing stations
Rockland	• None reported
Seattle	• Grades 7, 8, 9, reading and mathematics: Discipline problems; disruption of one testing period by a false fire alarm; confusion during testing
Selmer	• None reported
Taft	• None reported
Wichita	• Elementary students: restless, unmotivated students; extensive testing previous to project testing; both days of evaluation testing compressed into one day; unskilled test administrator for two third-grade control groups (36 students total) • Secondary students: unruly, disinterested, unmotivated; some students marking answers at random

For five other sites (Dallas, Hammond, Hartford, Philadelphia, Wichita), incidents and/or test conditions were not reported as specific to a grade, group, or subject matter level. Rather, certain general conditions were cited for either the elementary level, junior high level, or both levels. Thus, at Philadelphia, extreme heat (95 degrees) was reported during all testing sessions, and none of the testing rooms was air conditioned. At Dallas, certain organizational problems were cited at the junior high level, particularly difficulties in identifying the target population to be tested (several students were tested who were later identified as not being in the program, plus a high absence rate at the time of testing for those identified as being in the program). In turn, it was judged that these problems might have affected students' behavior during actual testing. How many students might have been affected is unknown.

In Hammond, Hartford, and Wichita, test coordinators reported problems concerning the maintenance of student discipline, and the test-taking motivation of students. At Hammond, this was reported for the junior high students. At Hartford, the coordinator reported that "testing sessions appeared to be conducted as efficiently as possible considering the type of youngster being evaluated; that is to say that the students in this project in this system were rather undisciplined, did not follow directions well, and had a tendency to put little effort into this total evaluation process". At Wichita, both elementary and secondary students were reported as being "turned-off" to the whole situation, as manifested in restless and sometimes unruly behavior, and manifested by some students at the secondary level sleeping through the test or marking answers on the tests at random. At the elementary level, contributing factors at Wichita were the reported extensive testing of students prior to the project testing, and the compressing of the testing into one day, rather than spreading the testing over two days as planned.

For the remaining six sites (Athens, Fresno, Jacksonville, Las Vegas, Portland, Seattle) incidents were reported specific to a given grade, group (E or C), and subject matter combination. In Table 17, for these six sites, the grade(s), group(s), and subject area(s) for which the incidents are reported are indicated first, followed by a brief description of the incident(s). Also given, when available, is an estimate of the percentage of students within a given grade(s), group(s), and subject matter area(s) for which the incident pertains.

As can be seen in the table, for these six sites, again problems centering on discipline and lack of test-taking motivation were the type of incidents reported, with these reports occurring at the junior high level. In addition, however, for certain of these sites, as shown by the table, conditions of overcrowding were reported, and also the testing of groups at a testing session in larger numbers than planned. In turn, these conditions probably accentuated to some extent the discipline problems and lack of student interest in taking the test. Also, at one site, "poor handling" of one of the grade 7 testing groups was a factor in the reported discipline problems. At another site, the testing of two or three grade levels (at the elementary) in a single room no doubt contributed to the problems and, at a third site, student behavior would probably have been improved had the testing effort been better planned at the site.

In summary, in those 13 sites where general or specific incidents or conditions were reported, the type of problem reported most often centered around discipline problems in the classroom and lack of test-taking motivation*. These problems were limited generally to the junior high level, and in several cases also limited to only some of the grades, schools, and testing groups within this level. In certain cases, these problems were accentuated by potentially "controllable" factors, such as overcrowding in the testing rooms, number of students tested at a sitting, amount of testing done in one day, etc. The latter implies that a retesting of certain sites might have yielded better testing conditions (as with Grand Rapids and Bronx). However, decisions to retest had to be based on several factors and considerations, not the least of which were budgetary considerations, judgments as to numbers of students possibly affected, and, quite importantly, for a given site the feasibility and acceptance by the site of further disrupting students' learning schedules by yet more testing. Moreover, it seems clear that the general problem centers more on the students who were tested, rather than on "test conditions", especially at the junior high level. The performance incentive experiment is concerned with students many of whom have lost faith in the school program and are simply not motivated to do well in most school functions, including their performance on standardized achievement tests. In turn, if youngsters were "turned-off" on a first testing, a second testing shortly following might have yielded even more student disinterest and behavioral problems.

---

* To a lesser degree, problems were created by administrative confusion resulting from the opening of school, late selection of students to be tested, and sometimes lack of communication to building personnel.

Moreover, several of the types of incidents and conditions reported for certain sites are not uncommon to any large scale standardized testing programs. The effect that such incidents have on test results, combined with the more unique problems in this study of student motivation and discipline for this target population, is not known. However, it is important that these circumstances be as explicitly noted as is possible, for they do serve to qualify the results and conclusions subsequently made in the last section of this report, at those sites, grades, and subject areas where incidents were reported.

Posttesting

Planning

Experience from the pretesting effort provided valuable background for facilitating an organized and smoothly running posttesting program. Advance planning by personnel at Battelle resulted in several procedure and scheduling changes that virtually eliminated the administrative, student identification, and fatigue problems. Changes which proved instrumental in overcoming some of the pretesting problems were

(1) A Battelle representative visited each site for a 1-4 day period approximately 1 month before testing was to begin. The purpose of this planning visit was to meet with the Project Director and his staff,

building principals, technology company representatives, teachers, and any other necessary parties to specify and agree upon schedules and responsibilities for the entire posttesting effort.

(2) Suggested testing schedules were altered to avoid testing on Fridays and Mondays. This gave experimental students a 4 to 5-day break between evaluation and certification testing and enabled larger numbers of students to be tested because Friday and Monday have proved to be high absentee days in most schools.

(3) Tests arrived on site with all student ID information already on the booklet. As a result, no confusion could arise as to which students were to be tested and which tests each student should take.

Implementation

In general, posttesting went extremely well in all the sites. The number of students tested was very high, and reports on testing conditions from the Battelle coordinators were favorable. Logistics posed few problems, the same high standards were met for recruiting qualified coordinators, materials were well prepared, and detailed plans for each site were available for the coordinator to study before he arrived on site.

Testing Conditions

As with the pretesting, Battelle coordinators were asked to report events which, in their judgment, might have endangered the validity of test results. These incidents are summarized in Table 18, which present incidents in the same format as given previously for the pretesting.

Although posttesting incidents centered around problems of discipline and test-taking motivation, there were somewhat fewer of these kinds of incidents, and they were generally less severe. In the three situations where problems were judged unusual, the testing was stopped and resumed only when conditions had improved.

In general, in context of the entire posttesting effort, those incidents that did occur were relatively small in number and often minor in nature, and tended to be localized to few testing groups. However, as with the pretest, circumstances and conditions noted in Table 18 serve to qualify conclusions and results in the analysis section of this report, at those grades, sites, and subject areas where incidents were reported.

TABLE 18. TESTING INCIDENTS REPORTED FOR POST EVALUATION TESTING

Site	Incidents or Conditions
Anchorage	• Grades 1 and 3, some experimental groups, reading and math: inadequate time, hurried directions, no breaks between tests.
Athens	• Grades 7, 8, 9, control: minor discipline problems, overcrowded cafeteria.
Bronx	• Grade 9, control, not tested • Grades 7, 8, 9, experimental and control: poor motivation, unruly students, partioned ballroom used for testing.
Dallas	• Grade 1, experimental, reading, one class (approx. 25% of students): hyperactive students during evaluation reading tests, regular teacher quieted them but minimal attention during rest of evaluation tests.
Fresno	• Grades 7, 8, 9, experimental group, reading: students refused to respond to substitute teachers as test administrators; and severe discipline problems were encountered; However, retesting was accomplished.
Grand Rapids	• None reported.
Hammond	• Grade 2, control, math: time limits for math subtests severely violated, for approximately 20% of students.

TABLE 18. (Continued)

Site	Incidents or Conditions
Hartford	• 7th grade, control, reading: schedule and time limits violated for one class (approx. 50% of students).
Jacksonville	• 7th grade, experimental, reading: crowded room, uncomfortable folding chairs, subpar handling by test administrators
Las Vegas	• None reported.
McComb	• None reported.
Philadelphia	• Grade 3, experimental, reading and math: outside noise, poor moral due to cancellation of field trip because of tests
Portland	• None reported.
Rockland	• None reported.
Seattle	• Grades 7, 8, 9, experimental: crowded rooms, disruptions during testing, poor discipline due to conflicts between company staff and regular school staff.
Selmer	• None reported.
Taft	• None reported.
Wichita	• Grade 2, experimental: one class extremely unruly due to lack of control by regular teacher. • Grades 7, 8, 9, experimental: several small groups, experienced discipline problems due to inability of proctors/examiners to maintain discipline.

THE DATA-ANALYSIS METHOD

The purposes of this section are to summarize the methodological issues studied during the selection of the method of analysis, to present alternative analysis methods examined, and to present a description of the method selected. A brief statement is given of each methodological issue studied, and its effect on the selection of the method of analysis is delineated. The description of the selected method includes a specification and discussion of the analysis models, and associated statistical tests employed to obtain the principal results of the analysis.

Issues in the Selection of an Analysis Method

A number of recognized methodological issues relevant to the analysis of the outcomes of this program were studied. Many of these issues arose from typical, unavoidable difficulties associated with "real world" implementations of experimental designs. Other issues were concerned with the identification of a suitable methodology for data analysis. There were three subject areas which were judged to include those issues relevant to the selection of an appropriate analysis method:

- Measurement of change
- Nonrandom sampling
- Measurement error.

A breif description and discussion of these areas follows.

Measurement of Change

Despite the fact that much controversial literature has been generated within the social sciences on the proper methods for measuring change (e.g., change from pretest to posttest), very little is of actual relevance to the problem of this experiment. For this experiment, the important, underlying question is not whether to measure "change" from pretest to posttest, per se, nor is it whether to use a "gain" or "change" score as the dependent variable. Rather, the basic question is how to take into account appropriately differences between the E and C groups on the pretest and other variables potentially related to posttest scores, when comparing the two groups on posttest performance. In the analyses described in this report, the pretest and other selected variables are taken into account by incorporating them into the analysis models as independent variables.

Nonrandom Sampling

Probably the greatest number of methodological issues resulted from the inability to obtain a random sample of students, schools, or groups from some well-defined population. Because of nonrandom sampling, an attempt was made to identify and measure several important variables upon which the experimental and control groups could be different. For example, several important student variables were measured, such as pretest score, race, and several student family background variables (see the section entitled "Description of Target Population"). However, other important student variables were not measured, such as student motivation. Also, the measurement and analysis of E and C school differences on such variables as teachers' attitudes toward the experimental program and the schools' curriculum outside of the experimental program were not within the scope of the study. Such differences in student and school variables offer competing explanations for the observed effects of the remedial program.

A partial solution to this problem was accomplished by means of quantitatively adjusting the observed effects of the remedial program on the basis of differences between the groups for some of the variables that were measured. The extent to which this was carried out appropriately is partly a function of the limited sample sizes at each grade/site/subject combination. Also, the choice of an appropriate model that includes several of these variables on which the groups might differ and that appropriately represents the functional relations between these variables is difficult -- and again, even more difficult due to limited sample sizes.

However, group differences on some of the more important variables were taken into account quantitatively in the analyses. Variables that were taken into account were pretest, student race, father's education, total family income, and an attitudinal item from the Parent Questionnaire concerning parents' approval of new teaching methods for their children. These particular variables were chosen because they were judged to have a potentially high relationship with posttest performance, because there were observed differences between the E and C groups on these variables at many grade/site combinations (see the section on "Description of Target Population"), and because these variables were judged to provide sufficient data within most grade/sites to warrant inclusion in any analysis. Father's occupation was considered also but was excluded due to the relatively low response rate to that item.

Differences on other variables measured in the study were taken into account only qualitatively. Because group differences were judged to exist on these other measured variables, and because group differences may exist on variables not measured, the conclusions presented are appropriately qualified.

Also, because nonrandom sampling creates uncertainty about the actual sampling distributions of commonly calculated statistics, such as "t", "F", and chi-square, the validity of statistical tests of significance is threatened. That is, the probability formally associated with any one of these statistics taking on an extreme value is actually unknown and can only be taken as an approximation. In the following analyses, the t-statistics calculated are only "formal" calculations. The probabilities associated with these calculations are assumed to be approximately correct. Usually, a t-value greater in absolute value than 2.0 is taken as evidence for a nonrandom result in all tests for significance. In some analyses, the limited sample sizes require a higher t-value in order to maintain the significance level at approximately 0.05.

Measurement Error

The assumption that independent variable values do not represent hypothesized "true" values raises the issue of "measurement error" in the analysis model. The presence of measurement error in such variables as pretest and posttest scores is usually indicated by estimating the reliability of such tests. The publishers' estimates of the reliability of the achievement tests used in this experiment are discussed in an earlier section. These estimates are consistently high. However, in a study dealing with a target population of low-achieving students from low-income families, reliability estimates based on samples from such a population are also important. Such estimates were obtained. Table 19 presents reliability estimates using an internal consistency measure (Kuder-Richardson 21). As can be seen from the table, all of these estimates are quite high, the lowest being 0.90.

These reliability estimates were obtained using the Kuder-Richardson 21 formula for the following reasons. An internal consistency estimate was judged most appropriate because the design of the study did not provide for obtaining additional test data at any other times than the beginning and end of the 1970-71 school year and the beginning of the 1971-72 school year. Thus, reliability estimates involving fluctuations in a student's score over a short time interval as a source of error, such as the correlation of alternate forms administered 2 weeks apart, were ruled out. The best estimate of reliability based upon an internal consistency measure is that most commonly referred to as the Kuder-Richardson 20 formula. All reliability estimates involving the correlation of variously defined halves of a

TABLE 19. ESTIMATED RELIABILITY COEFFICIENTS (KUDER-RICHARDSON 21) OF THE PRE- AND POSTTEST FORMS OF THE EVALUATION READING AND MATHEMATICS TEST FOR EACH GRADE(a)

	Reading		Mathematics	
	KR-21	N(b)	KR-21	N(b)
Grade 1				
Pretest (SEAT, Level I)(c)	0.94	2139	0.94	2124
Posttest (CAT, Level I, Form B)	0.90	2139	0.92	2124
Grade 2				
Pretest (MAT, Primary I, Form F)	0.92	2702	0.88	2531
Posttest (MAT, Primary I, Form G)	0.98	2702	0.97	2531
Grade 3				
Pretest (MAT, Primary II, Form F)	0.94	2482	0.98	2357
Posttest (MAT, Primary II, Form G)	0.96	2482	0.94	2357
Grade 7				
Pretest (MAT, Intermediate, Form F)	0.93	2319	0.93	2286
Posttest (MAT, Intermediate, Form G)	0.94	2319	0.95	2286
Grade 8				
Pretest (MAT, Advanced, Form F)	0.91	2256	0.90	2153
Posttest (MAT, Advanced, Form G)	0.92	2256	0.93	2153
Grade 9				
Pretest (MAT, Advanced, Form F)	0.93	2089	0.94	2077
Posttest (MAT, Advanced, Form G)	0.93	2089	0.94	2077

(a) The sample used to estimate KR-21 were full-year students with both a pre- and posttest score in the appropriate subject.

(b) N = the number of students in each sample.

(c) See an earlier section of this report for more complete identification and discussion of each test.

test, such as "split-half" and "odd-even", provide estimates of the value calculated using the Kuder-Richardson 20 formula. In fact, the average of all correlations between all possible halves of a test is equal to the Kuder-Richardson 20 value. However, the Kuder-Richardson 20 formula was not used to estimate reliabilities because it requires that a pass-fail response to each item for each student be stored in the data file. Such storage was not within the scope of this study. However, if one assumes that the item difficulty (percent who passed an item) of each item is equal to the average item difficulty, the Kuder-Richardson 20 formula reduces to a formula involving only the number of test items, and the mean and variance of the raw scores for the students in the sample. This formula is most commonly referred to as the Kuder-Richardson 21 formula. Finally, it is always true that this formula will provide a lower estimate of a test's reliability than the Kuder-Richardson 20 formula. This fact makes the high test-reliability estimates obtained in this study even more impressive.

In Battelle-Columbus' Interim Report on this experiment, qualitative judgments were made about the pretests' reliabilities in order to qualify any analysis results which might be suspect due to potentially unreliable pretest scores. These judgments were made on the basis of the correlations between pretest and posttest scores within the E and C groups at each site. At this time, however, the high estimated reliabilities in Table 19 are most encouraging and obviate the necessity for such judgments. It is still encouraging, however, to note that the large majority of these pre-post correlations are relatively high, as shown in Table 20. Table 20 provides a frequency distribution of the pre-post correlations*, by grade and subject. This table shows that the values of the pre-post correlations increase by grade level in the elementary grades and are generally high (above 0.70) in the junior high grades.

---

*See Appendix K for the values of the pre-post correlations within E and C groups, for each grade, site, and subject area.

TABLE 20. DISTRIBUTION OF PRE-POST CORRELATIONS FOR EXPERIMENTAL AND CONTROL GROUPS BY GRADE AND SUBJECT AREA

Value of r	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
	R	M	R	M	R	M	R	M	R	M	R	M
.00- .10	-	-	-	-	-	-	-	-	-	-	-	-
.11- .20	-	-	-	-	1	-	-	-	-	-	-	-
.21- .30	3	1	-	-	-	-	-	-	1	1	-	-
.31- .40	1	3	3	-	1	1	1	2	1	1	-	-
.41- .50	6	8	5	2	4	2	-	2	1	3	1	1
.51- .60	10	5	16	8	4	2	2	3	5	3	1	2
.61- .70	6	11	9	12	9	11	8	4	6	7	1	5
.71- .80	7	6	3	14	16	14	11	11	14	10	12	11
.81- .90	1	-	-	-	1	6	12	14	5	10	17	11
.91-1.00	-	-	-	-	-	-	2	-	3	1	2	4
Totals:	34	34	36	36	36	36	36	36	36	36	34	34

R = Reading
M = Mathematics

The correlations at Grade 1 are understandably lower because the pre- and posttests were not alternate forms of the same test. The pretest was a readiness test and the posttest was a standardized achievement test. In reading at Grades 2 and 3, most correlations are above 0.50 and 0.60, respectively. In mathematics at Grades 2 and 3, most are above 0.60 and 0.70, respectively. Overall, very few correlations are less than 0.40.

Selection of an Analysis Method

Five methods were considered as candidates for the analysis of the data in this study:

- compare the difference between the E and C posttest means with the difference between the E and C pretest means.

- assume a randomized blocks experimental design with pretest score as the blocking variable and analyze the results using the analysis of variance.
- assume a completely randomized experimental design with pretest score as a covariate and analyze the results using the analysis of covariance.
- compare the posttest scores for the E group with predicted scores for the E group obtained from a regression model fitted to the data for the C group.
- treat posttest score as a dependent variable; treat group membership, pretest scores, and selected other variables as independent variables, and analyze the results using regression analysis.

The first method was judged unacceptable because it does not provide a quantitative adjustment in mean posttest differences due to mean pretest differences and does not consider the selected other variables at all. The second method was not considered appropriate because the blocking variable is the pretest score and it may interact with the group variable. These conditions make "blocking" undesirable. The third method can account for a linear relationship between pretest and posttest, but it requires assumptions analogous to the block design that do not permit the study of interactions. The fourth method was judged unacceptable because it does not use the data from the E group to fit the regression model.

The fifth method can account for linear and nonlinear relations, and it allows the inclusions of interactions of pretest and other variables with the group membership variable. Group differences on the posttest may then be examined with the understanding that they have been adjusted for differences on the pretest and other variables and that any necessary interactions have been included in the analysis model. For these reasons, this method was selected for the analysis of the data.

Regression Analysis Models Employed

Two types of regression models were employed to provide the principal results. One involved group membership, pretest score, and their interaction as independent variables (called the Pre-Post Model). The second included pretest score and the family background and attitudinal variables discussed earlier without any interaction terms (called the Extended Variables Model).

Both of these models take into account only linear relations between pretest and posttest. Examination was made of the appropriateness of using only linear functions of pretest scores. This was accomplished by examining 236 computer-generated scatter plots. Each scatter plot represented each grade/site/subject combination, illustrating the actual relationship between pretest and posttest. Study of these plots indicated that nonlinear terms involving pretest were unnecessary. However, a few of the scatter plots at the elementary level did indicate some curvalinearity, in that there appeared to be, in these cases, a smaller rate of change (or slope) of posttest scores on pretest scores for the higher values of pretest. In most of these cases, the grade/sites involved were those where the pretest scores were overall relatively high and the higher scoring students on the pretest did not have as much room to improve (a ceiling effect) as the lower scoring students.

An example of a typical scatter plot for each grade is presented in Appendix L.

The Pre-Post Model

The data for each grade/site combination were fitted for the subjects Reading and Mathematics scores using a regression model of the following form:

$$Y_i = A + B_1X_{1i} + B_2X_{2i} + B_3X_{1i}X_{2i} + e_i$$

where A, B_1, B_2, B_3 denote regression coefficients and

Y_i = posttest score for student i, i = 1, ..., n;

X_{1i} = 1, if student i belonged to the experimental group, or

0, if student i belonged to the control group;

X_{2i} = pretest score for student i; and

e_i = deviation between the actual posttest score and the estimated posttest score using the fitted regression model.

Substitution of $X_{1i} = 1, 0$ into the model yields

$$Y_{1i} = (A + B_1) + (B_2 + B_3) X_{2i} + e_{1i}$$

and

$$Y_{0i} = A + B_2X_{2i} + e_{0i}$$

as the individual regression lines fitted to the experimental and control data, respectively. The Pre-Post Model was fitted to the data using the SPSS* regression program as implemented on Battelle-Columbus Laboratories' CDC 6400 computer. The difference between these regression lines after they have been estimated from the sample data** is then given by

$$\hat{Y}_1 - \hat{Y}_0 = \hat{B}_1 + \hat{B}_3X_2.$$

This expression shows that the difference between the regression estimate, $\hat{Y}_1$, of the posttest score for experimental students and the regression estimate, $\hat{Y}_0$, of the posttest score for control students is given by a linear function, $\hat{B}_1 + \hat{B}_3X_2$, where X_2 denotes a given specified pretest score. This is to say that the estimated difference between the two groups on the posttest varies for different values of pretest. Such variation is a linear function of the pretest score in this model.

Figure 2 provides a graphic illustration of this regression model. The difference between estimated posttest scores for the two groups is the difference between the two regression lines evaluated at any pretest level. Two such differences are described. One is the difference at the point where the pretest score is zero. This difference is algebraically equal to $\hat{B}_1$, the estimated coefficient of the group membership variable. Its magnitude and sign answer the question "What is the estimated difference between the groups when the pretest score is zero?" Such a question carries little substantive meaning. A more meaningful question is to ask what the difference between the groups is estimated to be for various values within the pretest score range for which data exist in both the E and C groups. An example of such a score range is illustrated in Figure 2 by the two vertical broken lines. Note that, because of the difference in slopes of the two lines (measured by $\hat{B}_3$), the difference between groups can change in magnitude and direction as the pretest level changes. For example, the group difference favors the C group when $X_2 = 20$, but favors the E group when $X_2 = 60$.

---

* Statistical Package for the Social Sciences, Nie, N., Bent, D.H., Hull, C. H., McGraw-Hill Book Company, New York, New York, 1970.

**The carat above each letter indicates that it has been estimated from sample data. After estimation, the subscript, i, is dropped.

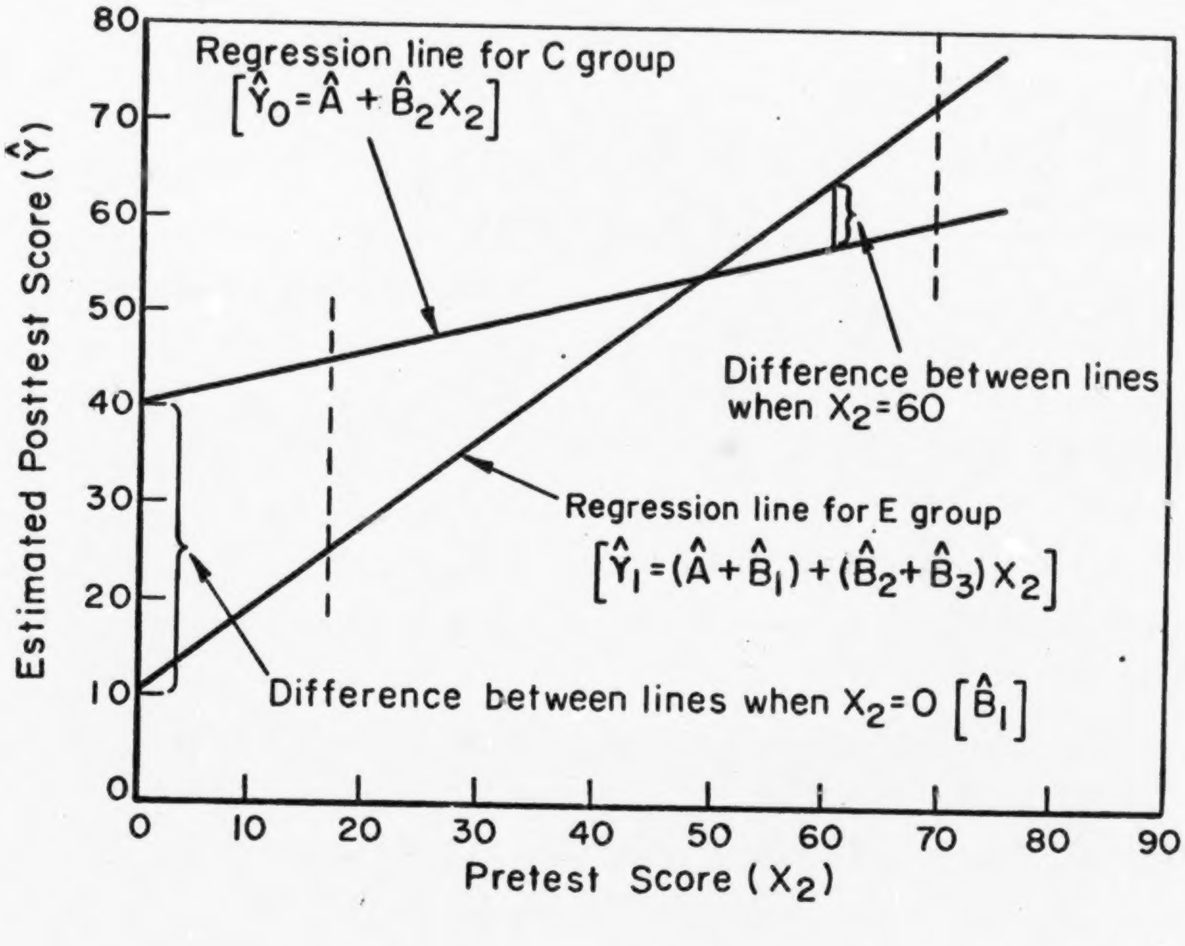


FIGURE 2. ILLUSTRATION OF THE "PRE-POST MODEL"

The Method Employed for Obtaining Principal Results With the Pre-Post Model. In order to assess the impact of the experimental program for any grade/site/subject, the difference, $\hat{Y}_1 - \hat{Y}_0$, was evaluated at the mean pretest, $\bar{X}_2$, taken over both the experimental and control students within each grade/site/subject combination. Each resulting difference was then divided by its estimated standard error, s_0, to obtain a ratio having the form of a t-statistic. Values of t greater than 2.0 were then taken as evidence for a significant group difference.

The standard error of the difference, s_0, was obtained as follows. Because $\hat{Y}_1 - \hat{Y}_0$ is a linear combination of $\hat{B}_1$ and $\hat{B}_3$ it may be shown that

$$s_0^2 = s_1^2 + \bar{X}_2^2 s_3^2 + 2\bar{X}_2 s_{13},$$

where

s_0^2 denotes the estimated variance of $\hat{Y}_1 - \hat{Y}_0$ evaluated at $X_2 = \bar{X}_2$;

s_1^2 and s_3^2 denote the estimated variances of $\hat{B}_1$ and $\hat{B}_3$, respectively;

and s_{13} denotes the covariance between $\hat{B}_1$ and $\hat{B}_3$.

Values for s_1^2 and s_3^2 were obtained from the computer printout of the regression results. Values for the covariance, s_{13}, were obtained using the formula:

$$s_{13} = \frac{s^2}{(nr_{13}s_{x_1}s_{x_3})} \left[\frac{r_{12}r_{23}r_{13} - r_{13}^2}{1 + 2r_{12}r_{23}r_{13} - r_{12}^2 - r_{23}^2 - r_{13}^2}\right],$$

where

s^2 denotes the residual mean square;

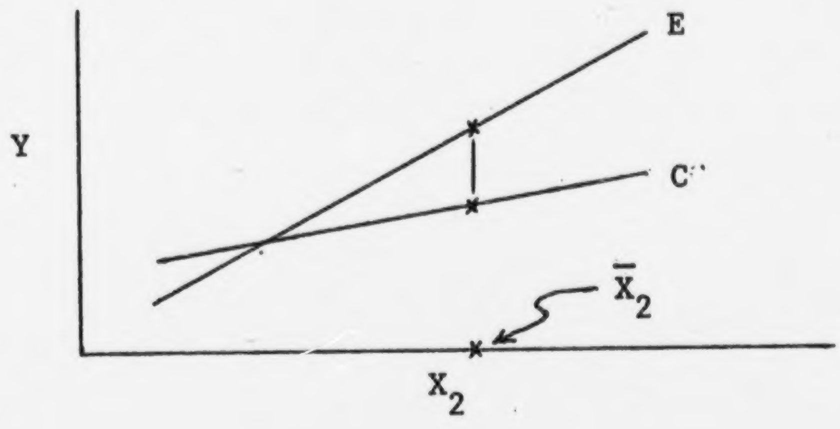
r_{12}, r_{13}, r_{23} denote the estimated correlation coefficients between group (X_1) and pretest (X_2), between X_1 and the group-by-pretest interaction (X_1X_2), and between X_2 and X_1X_2, respectively;

s_{x_1}, s_{x_3} denote the estimated standard deviations for X_1 and X_1X_2;

and n denotes the number of students.

This particular formula for calculating s_{13} was used because the values on the right-hand side were available from the computer printout. It was derived from the relation $s_{13} = c_{13}s^2$ where c_{13} denotes the (1,3) element of the inverse of the matrix of sums of squares and cross-products of deviations for X_1, X_2, and X_3.

To illustrate this t-test, one can draw the fitted regression lines for the E and C groups and evaluate them at the combined E and C mean pretest score. The difference between the lines, $\hat{Y}_1 - \hat{Y}_0$, at this point is the difference tested for significance. In the following sketch:



$\overline{X}_2$ represents the combined E and C mean pretest score. $\hat{Y}_1$, the estimated mean posttest score for the experimental program, is shown by the mark on the E line; $\hat{Y}_0$, the estimated mean posttest score for the traditional program, is shown by the mark on the C line; and the difference between the two estimated means $(\hat{Y}_1 - \hat{Y}_0)$ is shown by a vertical line between the two marks. This estimated posttest difference (at the combined E and C mean pretest score) forms the numerator of the t-statistic employed to test for significant group differences. This test forms the basis for the findings concerning experimental program effects in this study. If this t-test yields a value greater than +2.0, this can be taken as evidence that, in a larger collection of students (at the particular grade/site) with a pretest score distribution similar to the combined E and C pretest score distribution observed in the sample, the mean posttest score for those students if given the experimental program would be higher than the mean posttest score of those students if given the traditional program. Of course, if this t-test yields a value less than -2.0, one would predict a lower mean posttest score for students if given the experimental program than if given the traditional program.

In the above sense, the t-test employed provides an assessment of the "overall" effect of the E and C treatment conditions qualified, as mentioned earlier, on the basis of group differences on variables other than pretest. The outcome of this t-test, however, is not qualified if group differences exist on pretest. This is because the estimated means for the experimental and traditional programs are generated at a given pretest value, or for a given pretest distribution, so that any differences between the two groups in estimated posttest means cannot be attributed to group differences on pretest. In contrast, any difference between the E and C groups on their actual posttest means can be attributed partially to differences between the groups in their initial entry level. Thus, the regression technique, in the above outlined way, "takes into account" or "adjusts for" initial differences between the groups on the pretest. This concept is extended to adjust for group differences on variables in addition to pretest, in the Extended Variables Model, discussed subsequently in this section.

Study of the Significant Group-by-Pretest Interactions. The "Pre-Post Model" fitted to the data at each grade/site/subject combination includes a group-by-pretest interaction in the slopes of the E and C posttest regression lines as functions of pretest scores. If such a difference in slopes exists, then the inclusion of this term will result in an improved estimate of the residual error. This improved estimate, in turn, will yield a better statistical assessment of the difference between the E and C regression lines evaluated at the mean of the pretest scores for the combined E and C groups. This latter assessment is taken to be the primary measure of the impact of the experiment at each grade/site/subject combination, as mentioned above.

The inclusion of the interaction terms in the model also yields secondary benefits. Whenever the interaction is significant, the regression lines will have markedly different slopes, and may even intersect in the range of the pretest scores. This has the effect of showing reversed differences between the fitted E and C posttest means corresponding to low and high pretest scores. Because of the importance of this kind of information, a summary of the results obtained at all those grade/site/subject combinations which yielded a group-by-pretest interaction coefficient having a t-value greater than 2.0 is presented in Appendix M.

The Extended Variables Model

The data for each grade/site combination were also fitted, wherever possible, using a regression model with linear terms of the form:

$$Y_i = A + B_1X_{1i} + B_2X_{2i} + B_3X_{3i} \ . \ . \ . + B_pX_{pi} + e_i$$

where $A, B_1, \cdot \cdot \cdot, B_p$ denote the regression coefficients and

Y_i = posttest score for student i, $i = 1, \cdot \cdot \cdot, n$;

X_{1i} = 1, if student i belonged to the experimental group, or 0, if student i belonged to the control group;

X_{2i} = pretest score for student i;

$X_{3i}, \cdot\cdot\cdot, X_{pi}$ = total family income, father's education,* race, and parent's response to the "Approve New Method" attitude item, for student i; and

e_i = deviation between the actual posttest score and the estimated posttest score using the fitted regression model.

As in the Pre-Post Model, evaluating $X_{1i} = 1$ yields the model for the experimental group and $X_{1i} = 0$ yields the model for the control group. The difference between these two regression models after they have been estimated from the sample data is then simply

$$\hat{Y}_1 - \hat{Y}_0 = \hat{B}_1 \ .$$

---

* Father's education is divided into three levels (as defined in the section on "Description of Target Population") and, thus, requires two zero-one "dummy" variables. Race requires either none, one, or two zero-one "dummy" variables, depending upon whether a grade/site has one, two, or three predominant races.

That is, $\hat{B}_1$ is the estimated E/C difference in posttest means for any given value of pretest, or for that matter, for any given combination of values of the independent variables including the family background variables and the attitudinal items. Thus $\hat{B}_1$ is an estimate of experimental program impact holding constant the other variables included in the model, so that the value of $\hat{B}_1$ (the estimated difference in posttest means between experimental and traditional instruction) cannot be attributed to group differences in these other variables.

The group-by-pretest interaction term was not included in the Extended Variables Model because E/C group differences estimated with the Pre-Post Model were found to be quite similar (both in magnitude and statistical significance) to such differences in supplementary regression analysis performed at all grade/sites using a model with only group and pretest as independent variables. Overall, more than 90 percent of the statistical outcomes of the Pre-Post Model, comparing the E and C groups, remained unchanged when the interaction term was dropped in the supplementary analyses. This is not surprising when one considers that despite the presence of an interaction in the Pre-Post Model, the group comparison was made at the combined E and C pretest mean. Thus, for example, using the Pre-Post Model with data that demonstrated a strong cross-over effect (one group higher than the other at one end of the pretest score range and the opposite group higher at the other end; see Figure 2), the t-test made at the combined pretest mean would likely indicate a small, insignificant difference between groups. Using a model without the interaction term to analyze the same data would also likely yield a small, insignificant difference between groups over the entire pretest score range because the actual changing group differences over this range would tend to cancel each other out.

The Method Employed for Obtaining Results With the Extended Variables Model. In order to assess the impact of the experimental program for any grade/site/subject, the group coefficient, $\hat{B}_1$, is tested for significance using a t-test of the null hypothesis: $B_1 = 0$, at the 0.05 level of significance. Usually this means values of t greater than 2.0 are taken as evidence for a significant group difference. In some cases a value of t greater than 2.0 is required as the critical value, depending on the sample size and the number of variables in the model.

Summary

In summary, the analysis method used is regression analysis. Regression models are fitted at all grade/site combinations for both reading and mathematics. The Pre-Post Model includes pretest scores, group membership, and a group-pretest interaction as independent variables. The Extended Variables Model includes pretest scores, group membership, total family income, race, father's education, and an attitudinal item assessing attitudes of parents toward their children being taught by a new teaching method. The Pre-Post Model is also applied in additional analyses comparing the experimental group with the comparison group and the special treatment groups, comparing the experimental and control groups on attendance, and comparing the experimental and control groups on second year test scores obtained at the beginning of the 1971-72 school year. All of these applications are at only certain grade/sites. These additional analyses are discussed in greater detail in the following section on results.

RESULTS AND CONCLUSIONS: TECHNOLOGY COMPANY SITES

Introduction

The purpose of this section is to present the principal results obtained from a variety of analyses performed on the data. These results are presented under several headings. The main analyses are presented first. In general, the main analyses consist of comparing experimental and control groups at each site/grade/subject combination, utilizing in regression models pre- and post achievement test scores and data collected on other variables during the experimental (1970-71) year. The first part of these main analyses is the application of the Pre-Post Model, where E/C comparisons are made by regressing posttest scores against pretest scores. The resulting regression lines are then evaluated at the pretest mean of the combined experimental and control groups. The difference between these regression lines are tested for statistical significance using a formal t-value equal to 2.0. These results constitute the primary results of the Pre-Post Model analyses, and are summarized in Table 27.

The second part of the main analyses consists of applying, again at each site/grade/subject combination, extended regression models which include family background and attitudinal variables in addition to pretest. Results from these extended regression models are summarized in Table 30. The additional variables included in these extended models are family income, father's education, race of student, and one attitudinal item from the Parent Questionnaire. The attitudinal item included assesses parents' attitudes toward their children being taught by a new teaching method. As discussed earlier in the Target Population Section, E/C group differences existed very frequently on this item, and for this reason the item was included as a variable in the extended regression models.

The main analyses at each site are followed by an aggregate analysis across sites. This analysis is descriptive in nature, not involving any regression or other formal statistical models, and consists of an examination of pre- and posttest means for the experimental and control groups at each grade and in each subject area, aggregated across sites.

The changes in the grade equivalents of these means are then discussed.

In the next part of this section, results from analyses concerned with assesing the stability of the treatment impact are presented. To make this assessment, experimental and control students from 46 selected site/grade/subject combinations were administered the evaluation test during the fall of the school year that followed the experimental year. Those grade/site/subject combinations showing positive (or negative) impacts in the spring are examined to determine whether the impact remained positive (or negative) into the following year. These results are discussed and then summarized in Table 34. Because the original samples of both the experimental and control students suffered normal attrition during the summer months the pretest means for the reduced samples were computed and compared with those of the original samples. A discussion of these findings concludes the assessment of the stability of treatment impact.

As part of the experimental design for this program, many sites had "replacement" groups. These groups of students were generally taught in the same building as the treatment students. They were pretested and posttested as though they were control students. In some instances, as intended, replacement students were transferred out of their replacement groups and became members of experimental treatment groups whenever students were lost from the experimental treatment groups because of dropouts, etc. Where such transfers were few, the replacement groups were not severely depleted and became eligible for consideration as a kind of "in-house" control group (referred to in the analysis as "comparison" groups). This is in contrast to the normal control groups that were taught in other buildings. The posttest performance of students in 58 comparison groups are analyzed using the Pre-Post Model. Two regression analyses involving comparison groups are made. The first of these consists of comparing the posttest performance of students in the comparison groups with that of the control groups; the second consists of comparing the posttest performance of students in the comparison groups with that of the experimental groups. The results of these analyses are summarized in Tables 37 and 38. Finally, these two regression analyses are interrelated as shown in Table 39.

In addition to the experimental groups under the guidance of the technology companies, the experimental design also included some "special treatment" groups at Grand Rapids, Michigan, and Hartford, Connecticut. These groups were part of other remedial education programs or projects in these two school systems. The posttest performance of students in these special treatment groups is compared with experimental and control groups at these sites. The results of these analyses are given in Tables 41 through 45.

Because of the possible importance of student attendance in the assessment of program impact, a special analysis of attendance was made which compared the experimental and control groups. This analysis is summarized in Table 47.

The section concludes with a summary of results and conclusions.

The Main Analyses: Regression Analyses of Experimental verses Control Groups on First Year Test Results

The purpose of this section is to present the principal results based upon tests of significance of the difference between the experimental and control groups' estimated posttest means, using the two types of regression models: the Pre-Post Model and the Extended Variables Model. It is important to emphasize that results and comparisons are given in terms of estimated posttest means and not the actual posttest means. Estimated posttest means are generated from the regression models fit to the data. As discussed previously in the Method of Analysis section, any comparison of actual posttest means suffers from the fact that E and C groups differed on the pretest, so that any group differences between actual means on the posttest can be partly attributed to differences between the groups in their initial entry level, as well as partly attributed to differences between the groups on other variables, such as family background variables. The regression analysis technique copes with this problem by providing estimated posttest means for the E and C groups at a given common pretest value, so that any difference in the E and C estimated posttest means cannot be attributed to original differences between the groups on the pretest. For the Extended Variables Model, where group differences in other variables in addition to pretest are taken into account, this concept is extended so as to provide estimated means and differences between estimated means "holding constant" or "partialing out" the variables on which the groups might differ.

However, actual posttest means are provided in Appendix K for the interested reader, for experimental and control groups, for each site in each grade and subject area. Corresponding grade equivalent values are also given (shown in parentheses beside the raw score mean). Also, pretest raw score means and associated grade equivalents are given in Appendix K, along with standard deviations, sample sizes, and pre-post correlation coefficients.

The students in the regression analyses given below are full-year students, i.e., those identified as being in either the experimental or control group from a time within 3 weeks of the administration of the pretest to the time of the posttest administration. Further, each student was required to have both a pre- and posttest score in order to be included in the analysis. The numbers of such students in the E and C groups are given in Appendix N.*

Pre-Post Model Results

Pre-Post Model results are presented below for reading and for mathematics, within each grade. Tables and descriptive discussions show the sites for which statistically significant group differences were found, the estimated posttest means for each group, their differences, and the combined E and C pretest mean associated withthe posttest differences.**

All analyses are done with raw scores; however, the grade equivalents of the raw scores are presented also to facilitate the interpretation of the results through the use of a commonly used metric.

---

* Appendix N also provides the number of students in the comparison and special treatment groups analyses.

** Appendix R presents these data for all grade/site/subject combinations regardless of whether a statistically significant difference was or was not found. Also, the associated t-ratios are given in this appendix.

Table 21 shows Grade 1 results for reading and mathematics for those sites exhibiting a statistically significant group difference at the combined E and C pretest raw score mean.** The sites are shown in Column 1. Column 2 shows the stanine value corresponging to the combined E and C pretest raw score mean. The estimated posttest means shown in Columns 3 and 4 were obtained by evaluating the fitted E and C regression lines at the mean pretest value. The differences between these estimated posttest means are shown in the last column. The numbers in parentheses in the last three columns give the grade equivalents that correspond to the mean raw scores and their differences.

The table shows, for example, that the difference between the estimated raw score posttest means at Selmer is +13 and this difference corresponds to a grade equivalent difference of +0.7. This raw score difference is equal to the difference between 68 for the experimental posttest mean and 55 for the control posttest mean, as estimated by evaluating the fitted regression lines at the mean pretest value.

An examination of the results in this table shows that the largest positive differences in posttest raw score means for reading and mathematics are equal to 21 and 14, respectively, and both of these occur at Jacksonville. The largest positive differences in grade equivalents occur in reading at Selmer (+0.7) and in mathematics at Jacksonville (>+0.6). The largest negative difference in raw score posttest means occurs in reading at Wichita (-16). In terms of grade equivalents, the largest negative difference also occurs in reading at Wichita (-1.0). Overall, positive differences in posttest means occur in two sites in reading and in three sites in mathematics.

** Since Battelle's Interim report, the Grade 1 results have been re-analyzed in that scores of a zero posttest have been dropped from the analysis. Results from this analysis are in this report, and are taken to be definitive. The dropping of students with zero posttest scores resulted in three fewer impacts in favor of the E group than reported in the Interim report, and six more impacts in favor of the control group, in reading and mathematics combined.

TABLE 21. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS IN READING AND MATHEMATICS AND DIFFERENCES BETWEEN THESE MEANS FOR SITES EXHIBITING A SIGNIFICANT GROUP DIFFERENCE IN GRADE 1

READING

Site	Mean Pretest Value*	Estimated Posttest Means		E/C Difference
		E	C	
Selmer	(3)	68 (1.6)	55 (.7)	+13 (+.7)
Wichita	(3)	53 (.6)	69 (1.6)	-16 (-1.0)
Las Vegas	(2)	51 (<.6)	57 (.8)	-6 (<-.2)
Philadelphia	(2)	51 (<.6)	57 (.8)	-6 (<-.2)
Seattle	(4)	63 (1.2)	73 (1.8)	-10 (- .6)
Portland	(4)	62 (1.2)	71 (1.7)	-9 (-.5)
Jacksonville	(2)	57 (.8)	36 (<.6)	+21 (>+.2)
Hammond	(3)	62 (1.2)	73 (1.8)	-11 (-.6)

MATH

Site	Mean Pretest Value*	Estimated Posttest Means		E/C Difference
		E	C	
Dallas	(1)	34 (1.0)	29 (.6)	+5 (+.4)
Las Vegas	(2)	27 (<.6)	40 (1.3)	-13 (<-.7)
Fresno	(4)	33 (.9)	42 (1.5)	-9 (-.6)
Philadelphia	(2)	28 (<.6)	38 (1.2)	-10 (<-.6)
Grand Rapids	(2)	32 (.9)	41 (1.4)	-9 (-.5)
Hartford	(2)	30 (.7)	39 (1.3)	-9 (-.6)
McComb	(2)	44 (1.6)	36 (1.1)	+8 (+.5)
Portland	(4)	41 (1.4)	55 (2.0)	-14 (-.6)
Jacksonville	(2)	38 (1.2)	24 (<.6)	+14 (>+.6)

* The stanine value corresponding to the combined E and C pretest raw score mean.

Table 22 gives the results for Grade 2 for those sites showing a significant group difference at the combined raw score mean for the E and C groups. The mean pretest values in Column 2 are given in raw score units along with corresponding grade equivalents. The regression estimates of the posttest means and the differences between them for the E and C groups are given in Columns 3, 4, and 5. Raw score values and corresponding grade equivalents are given for the estimated posttest means.

The largest positive differences in the posttest means occur at Jacksonville (+6) for reading and at Dallas (+9) for mathematics. The largest negative differences in these means occur at Las Vegas for both reading (-14) and mathematics (-7). In reading the grade equivalent differences range between a maximum at Wichita (0.2) and a minimum at Las Vegas or Bronx (-0.3); in mathematics this range is associated with Jacksonville (+0.4) and Rockland (-0.5). The E group at Rockland in mathematics shows a larger grade equivalent gain than any other E group in the table (2.6-1.7=0.9). All posttest means have grade equivalents less than 2.9 ("average" grade-level position at the end of the school year is 3.0) with the exception of the control group at Rockland in mathematics (3.1). The table shows 3 positive and 7 negative differences in posttest means in reading; and 2 positive and 7 negative differences in mathematics.

Table 23 shows the regression results for Grade 3. The maximum and minimum differences in raw score units are at Selmer (+9) and Seattle (-11) for reading, Dallas (+20) and either Las Vegas or Hartford (-13) for mathematics. Selmer and Seattle also yield the maximum and minimum differences in grade equivalents for reading, (0.4) and (-0.5), respectively. In mathematics these extremes are associated with Dallas (+0.6) and either Hartford or McComb (-0.4). It should be noted that a gain of 1.3 grade equivalents occurred for the E group in mathematics in Selmer. The results show 3 positive and 4 negative differences in the posttest means for reading, and 5 positive and 8 negative differences in the posttest means for mathematics.

TABLE 22. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS IN READING AND MATHEMATICS AND DIFFERENCES BETWEEN THESE MEANS FOR SITES EXHIBITING A SIGNIFICANT GROUP DIFFERENCE IN GRADE 2

READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Selmer	40 (1.7)	65 (2.2)	60 (2.1)	+5 (+.1)
Wichita	39 (1.7)	63 (2.2)	58 (2.0)	+5 (+.2)
Las Vegas	23 (1.3)	32 (1.5)	46 (1.8)	-14 (-.3)
Taft	37 (1.6)	52 (1.9)	57 (2.0)	-5 (-.1)
Hartford	27 (1.4)	42 (1.7)	51 (1.9)	-9 (-.2)
McComb	29 (1.4)	47 (1.8)	55 (2.0)	-8 (-.2)
Seattle	39 (1.7)	58 (2.0)	65 (2.2)	-7 (-.2)
Portland	42 (1.8)	59 (2.1)	65 (2.2)	-6 (-.1)
Jacksonville	27 (1.4)	50 (1.9)	44 (1.8)	+6 (+.1)
Bronx	28 (1.4)	41 (1.7)	54 (2.0)	-13 (-.3)

MATH

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Dallas	21 (1.2)	34 (1.5)	25 (1.3)	+9 (+.2)
Rockland	38 (1.7)	53 (2.6)	56 (3.1)	-3 (-.5)
Las Vegas	20 (1.2)	29 (1.4)	36 (1.6)	-7 (-.2)
Fresno	31 (1.4)	43 (2.0)	47 (2.2)	-4 (-.2)
Philadelphia	23 (1.3)	36 (1.6)	40 (1.8)	-4 (-.2)
Seattle	32 (1.5)	45 (2.1)	51 (2.4)	-6 (-.3)
Portland	39 (1.7)	49 (2.3)	52 (2.5)	-3 (-.2)
Jacksonville	23 (1.2)	44 (2.0)	36 (1.6)	+8 (+.4)
Bronx	22 (1.2)	40 (1.8)	43 (2.0)	-3 (-.2)

TABLE 23. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS IN READING AND MATHEMATICS AND DIFFERENCES BETWEEN THESE MEANS FOR SITES EXHIBITING A SIGNIFICANT GROUP DIFFERENCE IN GRADE 3

READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Selmer	45 (2.4)	67 (3.2)	58 (2.8)	+9 (+.4)
Dallas	26 (1.8)	36 (2.2)	29 (1.9)	+7 (+.3)
Las Vegas	28 (1.9)	39 (2.3)	47 (2.4)	-8 (-.1)
Seattle	46 (2.4)	53 (2.6)	64 (3.1)	-11 (-.5)
Jacksonville	30 (2.0)	44 (2.4)	38 (2.2)	+6 (+.2)
Hammond	42 (2.3)	54 (2.7)	59 (2.8)	-5 (-.1)
Bronx	31 (2.0)	47 (2.4)	53 (2.6)	-6 (-.2)

MATH

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Selmer	59 (2.4)	88 (3.7)	75 (3.2)	+13 (+.5)
Athens	50 (2.3)	68 (2.9)	63 (2.6)	+5 (+.3)
Wichita	39 (1.9)	53 (2.3)	60 (2.5)	-7 (-.2)
Dallas	35 (1.6)	59 (2.5)	39 (1.9)	+20 (+.6)
Las Vegas	37 (1.7)	49 (2.3)	62 (2.6)	-13 (-.3)
Fresno	51 (2.3)	62 (2.6)	68 (2.9)	-6 (-.3)
Taft	52 (2.3)	71 (3.0)	65 (2.7)	+6 (+.3)
Hartford	44 (2.2)	51 (2.3)	64 (2.7)	-13 (-.4)
McComb	44 (2.2)	62 (2.6)	70 (3.0)	-8 (-.4)
Seattle	53 (2.3)	71 (3.0)	78 (3.2)	-7 (-.2)
Portland	58 (2.4)	74 (3.1)	80 (3.3)	-6 (-.2)
Jacksonville	38 (1.8)	58 (2.4)	51 (2.3)	+7 (+.1)
Hammond	53 (2.3)	67 (2.8)	71 (3.0)	-4 (-.2)

Table 24 gives the results from the regression analyses for Grade 7. Only 2 sites in reading, Athens and Rockland, and 4 sites in mathematics, Rockland, Fresno, Grand Rapids, and Seattle, yielded significant group differences. The only positive difference between the posttest means occurred at Athens in reading. The table also shows that the larger grade equivalent gains were in two C groups in mathematics (1.2 for Rockland and 1.3 for Seattle). In reading there was one positive and one negative difference in posttest means. In mathematics all 4 differences were negative.

Table 25 shows the results obtained for Grade 8. In reading, the largest positive difference between the regression estimates of the posttest means occurred at Grand Rapids (+7); the largest negative difference occurred at Seattle (-10). In mathematics, there are no positive differences and the largest negative difference occurs at Seattle (-20). A large grade equivalent gain of 1.4 occurred for the E groups in reading at Anchorage and Grand Rapids; similar gains of 1.3 occurred for the C groups in Rockland, Seattle, and Hammond. A negative gain (-0.6) in grade equivalents occurred for reading in the E group of Philadelphia. In mathematics, the table shows that the grade equivalent gain for the C group at every site exceeds 1.0 with the exception of Hammond (0.9). The largest negative difference occurred for Seattle, where the E group showed no gain in grade equivalents and the C group showed a gain of 1.6 in grade equivalents. Three sites showed positive impacts in reading; no sites showed positive impacts in mathematics.

Table 26 shows the results obtained for Grade 9. In both reading and mathematics, 3 sites showed positive differences in the regression estimates of the posttest means. In reading, the largest gains in grade equivalents were made by the E group at Athens (1.2) and by the C group at Seattle (1.1). In mathematics, both the minimum and maximum gains in grade equivalents occurred at Seattle where the E group showed a grade equivalent gain of 0.6 and the C group showed a grade equivalent gain of 1.5.

TABLE 24. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS IN READING AND MATHEMATICS AND DIFFERENCES BETWEEN THESE MEANS FOR SITES EXHIBITING A SIGNIFICANT GROUP DIFFERENCE IN GRADE 7

READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Athens	38 (4.4)	43 (4.8)	39 (4.5)	+4 (+.3)
Rockland	56 (5.8)	59 (6.0)	64 (6.6)	-5 (-.6)

MATH

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Rockland	62 (5.7)	72 (6.4)	79 (6.9)	-7 (-.5)
Fresno	36 (4.3)	42 (4.6)	52 (5.2)	-10 (-.6)
Grand Rapids	38 (4.4)	44 (4.7)	50 (5.1)	-6 (-.4)
Seattle	57 (5.4)	59 (5.6)	75 (6.7)	-16 (-1.1)

TABLE 25. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS IN READING AND MATHEMATICS AND DIFFERENCES BETWEEN THESE MEANS FOR SITES EXHIBITING A SIGNIFICANT GROUP DIFFERENCE IN GRADE 8

READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Dallas	23 (3.7)	31 (4.6)	27 (4.2)	+4 (+.4)
Anchorage	46 (6.6)	59 (8.0)	53 (7.3)	+6 (+.7)
Rockland	47 (6.7)	55 (7.4)	59 (8.0)	-4 (-.6)
Fresno	30 (4.5)	33 (4.9)	38 (5.6)	-5 (-.7)
Philadelphia	26 (4.1)	22 (3.5)	31 (4.6)	-9 (-1.1)
Grand Rapids	32 (4.8)	42 (6.2)	35 (5.2)	+7 (+1.0)
Seattle	41 (6.0)	43 (6.2)	53 (7.3)	-10 (-1.1)
Hammond	40 (5.8)	47 (6.7)	51 (7.1)	-4 (-.4)

MATH

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Rockland	56 (6.8)	66 (7.8)	70 (8.1)	-4 (-.3)
Fresno	34 (4.8)	37 (5.4)	42 (5.9)	-5 (-.5)
Grand Rapids	38 (5.3)	39 (5.7)	47 (6.4)	-8 (-.7)
Seattle	47 (6.0)	44 (6.0)	64 (7.6)	-20 (-1.6)
Portland	47 (6.0)	52 (6.7)	60 (7.3)	-8 (-.6)
Hammond	46 (6.0)	51 (6.7)	55 (6.9)	-4 (-.2)

TABLE 26. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS IN READING AND MATHEMATICS AND DIFFERENCES BETWEEN THESE MEANS FOR SITES EXHIBITING A SIGNIFICANT GROUP DIFFERENCE IN GRADE 9

READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Athens	34 (5.0)	43 (6.2)	39 (5.7)	+4 (+.5)
Las Vegas	49 (6.9)	58 (8.0)	52 (7.3)	+6 (+.7)
Philadelphia	23 (3.7)	24 (3.8)	27 (4.2)	-3 (-.4)
Taft	36 (5.3)	45 (6.4)	37 (5.5)	+8 (+.9)
Seattle	49 (6.9)	54 (7.4)	58 (8.0)	-4 (-.6)
Portland	47 (6.7)	53 (7.3)	56 (7.6)	-3 (-.3)

MATH

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
		E	C	
Selmer	59 (6.9)	68 (8.0)	73 (8.4)	-5 (-.4)
Athens	41 (5.6)	53 (6.8)	48 (6.4)	+5 (+.4)
Anchorage	53 (6.6)	65 (7.8)	55 (6.9)	+10 (+.9)
Rockland	70 (7.8)	78 (9.0)	70 (8.1)	+8 (+.9)
Seattle	54 (6.7)	60 (7.3)	71 (8.2)	-11 (-.9)
Portland	54 (6.7)	61 (7.5)	67 (8.0)	-6 (-.5)

Summary of Pre-Post Model Results

Table 27 shows a summary of the results obtained with the pre-post regression model for each grade, site, and subject. The symbols, E and C, indicate the occurrence of a significant positive or negative group difference, respectively, between the posttest means as estimated by the regression lines fitted to the pre-post data. The marginal row and column totals show the number of E's and C's together with the number of empty cells where a nonsignificant (Not Sig.) group difference was obtained.

The table shows, for example, in Grade 2-reading, the positive impacts (E) occurred at Selmer, Wichita, and Jacksonville; the negative impacts (C) occurred at Las Vegas, Taft, Hartford, McComb, Seattle, Portland, and Bronx. The marginal totals for this column show 3 positive impacts; 7 negative impacts; and 8 sites where the group differences were not significant. An examination of the row for Wichita shows a negative impact in Grade 1-reading; a positive impact in Grade 2-reading; a negative impact in Grade 3-mathematics; and no significant impact in any other grade/subject combination. The marginal totals for Wichita show one positive impact, 2 negative impacts; and 9 instances where the impacts were not significant.

An examination of the column totals shows that five positive impacts were obtained for Grade 3 in mathematics. No other grade/subject combination shows a greater number of positive impacts. The smallest number of positive impacts for grade/subject combinations occurred for Grades 7 and 8 in mathematics where no positive impacts were obtained. The corresponding maximum and minimum number of negative impacts are seen to be 8 for Grade 3-mathematics, and one for Grade 7-reading.

In none of the 12 grade/subject combinations did the number of positive impacts exceed the number of negative impacts.

An examination of the row totals shows that Jacksonville exhibited six positive impacts. The seven sites showing no positive impacts are Fresno, Philadelphia, Hartford, Seattle, Portland, Hammond, and Bronx. The maximum number of negative impacts is shown by the ten

TABLE 27. SUMMARY OF PRE-POST MODEL RESULTS FOR EACH GRADE, SITE, AND SUBJECT

Site	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9		No. E's	No. C's	No. Not Sig.
	Reading	Math.	Reading	Math.	Reading	Math.	Reading	Math.	Reading	Math.	Reading	Math.			
Selmer	E		E		E	E						C	4	1	7
Athens						E	E				E	E	4	0	8
Wichita	C		E			C							1	2	9
Dallas		E		E	E	E			E				5	0	7
Anchorage									E			E	2	0	10
Rockland				C			C	C	C	C		E	1	5	6
Las Vegas	C	C	C	C	C	C					E		1	6	5
Fresno		C		C		C		C	C	C			0	6	6
Philadelphia	C	C		C					C		C		0	5	7
Taft			C			E					E		2	1	9
Grand Rapids		C						C	E	C			1	3	8
Hartford		C	C			C							0	3	9
McComb		E	C			C							1	2	9
Seattle	C		C	C	C	C		C	C	C	C	C	0	10	2
Portland	C	C	C	C		C				C	C	C	0	8	4
Jacksonville	E	E	E	E	E	E							6	0	6
Hammond	C				C	C			C	C			0	5	7
Bronx	No Data	No Data	C	C	C						No Data	No Data	0	3	5
No. E's:	2	3	3	2	3	5	1	0	3	0	3	3	28		
No. C's	6	6	7	7	4	8	1	4	5	6	3	3		60	
No. Not Sig.	9	8	8	9	11	5	16	14	10	12	11	11			124
	17	17	18	18	18	18	18	18	18	18	17	17			

negative impacts at Seattle. No negative impacts occurred at Athens, Dallas, Anchorage, and Jacksonville. The difference between the number of positive impacts and the number of negative impacts is seen to be a maximum of 6 for Jacksonville and a minimum of -10 for Seattle.

The lower right-hand totals show that 28 positive impacts occurred, as compared to 60 negative impacts. Thus, the control groups exhibited statistically superior performance as determined by the regression analyses in approximately twice or many instances or did the experimental groups. However, by far the major outcome was no statistically significant difference between the groups, since 124 differences were not significant.

Of the 28 positive impacts, almost twice as many (18) occurred at the Elementary level than at the Secondary level (10). However, at either level the percentage of positive impacts out of all grade/site/subject combinations is very low, being 17 and 9 percent respectively. It is also interesting to note that 23 of the 28 positive impacts occurred at sites associated with three of the six technology companies, as shown in Table 28. However, such a comparison needs to be interpreted with caution, since different companies did not conduct programs in the same sites. Therefore, it is impossible to disentangle the site from the company effects.

Extended Variables Model Results

The purpose of this section is to present the results of comparisons between the posttest performance of the experimental (E) and control (C) groups taking into account group differences on selected variables in addition to the pretest score. These variables are taken into account quantitatively through inclusion in the extended regression model described earlier. In review, this model is an extension of the Pre-Post Model without the interaction term between pretest and group. The additional variables selected were student race, father's education, total family income, and level of parents' approval of new instructional methods for their children (Item No. 10 on the Parent Question and referred to as "Approve New Method").

TABLE 28. NUMBER OF POSITIVE IMPACTS FROM THE PRE-POST MODEL ANALYSES FOR THE ELEMENTARY AND JUNIOR HIGH GRADES, FOR EACH TECHNOLOGY COMPANY

Company	Elementary	Secondary	Totals
A	6	3	9
B	4	4	8
C	0	1	1
D	1	2	3
E	1	0	1
F	6	0	6
Totals:	18	10	28

A:	Selmer, Athens, Wichita
B:	Dallas, Anchorage, Rockland
C:	Las Vegas, Fresno, Philadelphia
D:	Taft, Grand Rapids, Hartford
E:	McComb, Seattle, Portland
F:	Jacksonville, Hammond, Bronx

The rationale for selecting these variables and dropping the group-by-pretest interaction is discussed in the previous section on the methods of analysis. The results of the Extended Variables Model are presented here.

There are certain disadvantages in using the Extended Variables Model to assess program impact. A smaller data base is available for this model than with the Pre-Post Model because it contains only students whose parents returned a questionnaire. Not only are the samples smaller in size (generally 40-50 percent as large as the samples in the Pre-Post Model analyses), but also the samples could be different in the type of students in them. However, a comparison of combined E and C pretest means of the samples associated with the Pre-Post and Extended Variables Models showed that these pretest means were close, most often not differing by more than one or two raw score units. There were, however, exceptions, particularly at Hartford and Grand Rapids, where the corresponding pretest means often differed on the order of four to six raw score units. In general, when the means of the samples associated with the two models did differ, the sample associated with the Extended Variables Model (EVM) had a higher pretest mean.

Although there are the above disadvantages, by use of this model group differences on variables in addition to pretest can be taken into account. Thus, in this sense, the results of these extended model analyses do provide a better comparison of experimental and control groups.

Table 29 describes the additional variables included in the extended variables model for each grade/site. Note that no EVM was fitted at Rockland, Philadelphia, and Grades 1, 7, 8, and 9 at Bronx.

TABLE 29. INDEPENDENT VARIABLES CONSTITUTING THE EXTENDED VARIABLES MODEL FOR EACH GRADE/SITE COMBINATION*

		Independent Variables					
Site	Grade	Group	Pretest	Student's Race	Father's Education	Total Family Income	Approve New Method
Selmer	All	X	X		X	X	X
Athens	All	X	X	X	X	X	X
Wichita	All	X	X	X	X	X	X
Dallas	All	X	X		X	X	X
Anchorage	All	X	X	X	X	X	X
Rockland **	None						
Las Vegas	1	X	X	X			
Las Vegas	2-9	X	X	X	X	X	X
Fresno	All	X	X	X	X	X	X
Philadelphia **	None						
Taft	All	X	X		X	X	X
Grand Rapids	All	X	X	X	X	X	X
Hartford	All	X	X		X	X	X
McComb	All	X	X	X	X	X	X
Seattle	All	X	X	X	X	X	X
Portland	All	X	X		X	X	X
Jacksonville	All	X	X		X	X	X
Hammond	All	X	X	X	X	X	X
Bronx **	1&9						
Bronx	2&3	X	X	X	X	X	X
Bronx **	7&8						

* Presence of variable in EVM is indicated by an "X".

** No EVM's fitted due to insufficient or no data.

Table 30 presents the results of the Extended Variables Models (EVM) along with the results from the Pre-Post Model (PPM), so that comparisons of results from the two models can be made. For each grade/site/subject the following entry is made: "E" if the difference between groups was statistically significant in favor of the E group, a "C" if the difference was in favor of the C group, and a blank if the difference was not significant.

TABLE 30. COMPARISON OF SIGNIFICANT E/C GROUP DIFFERENCES FOUND USING THE PRE-POST MODEL (PPM) VERSUS THE EXTENDED VARIABLES MODEL (EVM) FOR EACH GRADE, SITE, AND SUBJECT

Site	Grade 1				Grade 2				Grade 3				Grade 7				Grade 8				Grade 9				No. E's		No. C's		No. Not Sig.	
	Reading		Math.		Reading		Math.		Reading		Math.		Reading		Math.		Reading		Math.		Reading		Math.							
	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM
Selmer	E	E			E	E			E	E	E	E											C		4	4	1	0	7	11
Athens											E	E	E			E					E		E		4	2	0	0	8	10
Wichita	C	C			E						C	C													1	0	2	2	9	10
Dallas			E				E	E	E		E	E					E					E			5	3	0	0	7	9
Anchorage																	E						E	E	2	1	0	0	10	11
Rockland	No Extended Variable Models																										-			
Las Vegas	C	C	C	C	C	C	C		C		C					C					E				1	0	6	4	5	8
Fresno			C	C			C	C			C	C			C	C	C		C						0	0	6	4	6	8
Philadelphia	No Extended Variable Models																													
Taft					C			C			E										E				2	0	1	1	9	11
Grand Rapids			C												C		E		C	C					1	0	3	1	8	11
Hartford			C		C	C		C			C									C					0	0	3	3	9	9
McComb			E		C						C			E		E									1	2	2	0	9	10
Seattle	C	C			C		C		C	C	C	C			C	C	C	C	C	C	C	C	C	C	0	0	10	8	2	4
Portland	C		C	C	C	C	C	C			C								C	C	C		C	C	0	0	8	5	4	7
Jacksonville	E	E	E	E	E		E		E		E	E													6	3	0	0	6	9
Hammond							C		C		C	C					C		C						0	0	5	1	7	11
Bronx	No Data				C				C				No Extended Variable Models								No Data				0	0	2	0	2	4
No. E's	2	2	3	1	3	1	2	1	3	1	4	4	1	1	0	2	3	0	0	0	3	1	2	1	26	15				
No. C's	4	3	5	3	7	3	5	4	4	1	8	4	0	0	3	1	3	1	5	4	2	1	3	2			49	27		
No. Not Sig.	9	10	7	11	6	12	9	11	9	14	4	8	14	14	12	12	9	14	10	11	10	13	10	12					109	142
	15	15	15	15	16	16	16	16	16	16	16	16	15	15	15	15	15	15	15	15	15	15	15	15						

The column totals give the number of "E's" and "C's" and blanks (Not. Sig.) for each grade/subject, summed across sites. For example, in Grade 8-Reading, of the 15 sites where both the EVM and PPM were employed, there were three "E's" and three "C's" as a result of using the PPM, but no "E's" and only one "C" based upon the EVM. This type of result--a loss in number of both "E's" and "C's"--is by far the most typical. Note, however, that in Grade 7-Mathematics there were no "E's" and three "C's" using the PPM, but two "E's" and only one "C" using the EVM.

The row totals show the number of "E's", "C's" and blanks for each site summed across grades and subjects for those sites where an Extended Variables Model was used. None of the sites gained in the number of "E's" or "C's" (with the exception of McComb, which gained one E). Of the four sites which had a relatively large number of "E's" using the PPM (Selmer, Athens, Dallas, and Jacksonville), only Selmer maintained the same number using the EVM, with Athens, Dallas, and Jacksonville losing in the number of "E's". In fact, of the 10 sites exhibiting at least one positive impact with the Pre-Post Model and where the Extended Variables Model was employed, in 8 of the 10 sites the number of positive impacts was reduced when using the EVM.

The totals in the lower right corner of the table describe the overall result. Where 26 "E's" were obtained with the PPM, only 15 were obtained with the EVM, in those grades and sites where both models were used. Forty-nine "C's" were found using the PPM, only 27 using the EVM. The corresponding number of nonsignificant outcomes went from 109 with the PPM to 142 with the EVM. The number of "E's" lost, 11, represents 42 percent of the number obtained with the PPM. The number of "C's", 22, represents 45 percent. Thus, the change to the EVM does not appear to have favored either group.

Although the change to the Extended Variables Model does not favor either group, the conclusion is nonetheless indicated that application of the Extended Variables Model provides even less evidence in favor of an experimental group impact than the Pre-Post Model, simply because the proportion of positive impacts is less. And, in general, those few sites exhibiting a relatively large number of positive impacts with the Pre-Post Model lost some of these positive impacts when the extended model was applied.

Experimental - Control Group Comparisons Aggregated Across Sites

Although the primary emphasis of the analyses of data in this report is on obtaining results at a grade/site level, by means of regression analyses, it is informative to make descriptive comparisons of the experimental and control gourps' average performance on the posttest aggregated over the 18 sites. Table 31 shows the raw score means (and associated grade equivalents)* of the pre- and posttests for the E and C groups at each grade for all 18 sites combined. As shown there, the differences in pre-post gains between the two groups, either in raw scores or grade equivalents, are very small. In terms of raw score units, the largest difference in gain between the two groups occurs in Grade 3 reading, where the experimental group gained 16 raw score points, as compared to a gain of 13 raw score points for the control group. However, in terms of grade equivalent units for Grade 3 reading, there is no difference in gains between E and C groups, each group gaining 0.4 grade equivalents.

The largest difference in gains between the two groups in terms of grade equivalent units occurs at Grade 3 mathematics, where the experimental group gained 0.4 grade equivalents, as compared to a gain of 0.7 grade equivalents for the control group. Note, however, that the difference in gain between the groups in raw score units is zero, each group in Grade 3 mathematics gaining 18 raw score units.

Figure 3 portrays the data in Table 31 graphically, by showing a plot of grade equivalent gains as a function of grade level. (No results are shown for Grade 1 because the pretest did not yield grade equivalents.) For mathematics, the upper plot shows, for example, that the third grade experimental group began the third grade with an average grade equivalent of 2.2, and ended the third grade with an average grade equivalent of 2.6.

---

* The grade equivalents were obtained by first calculating the raw score mean, and then determining the grade equivalent corresponding to this mean raw score value in the publisher's grade equivalence tables.

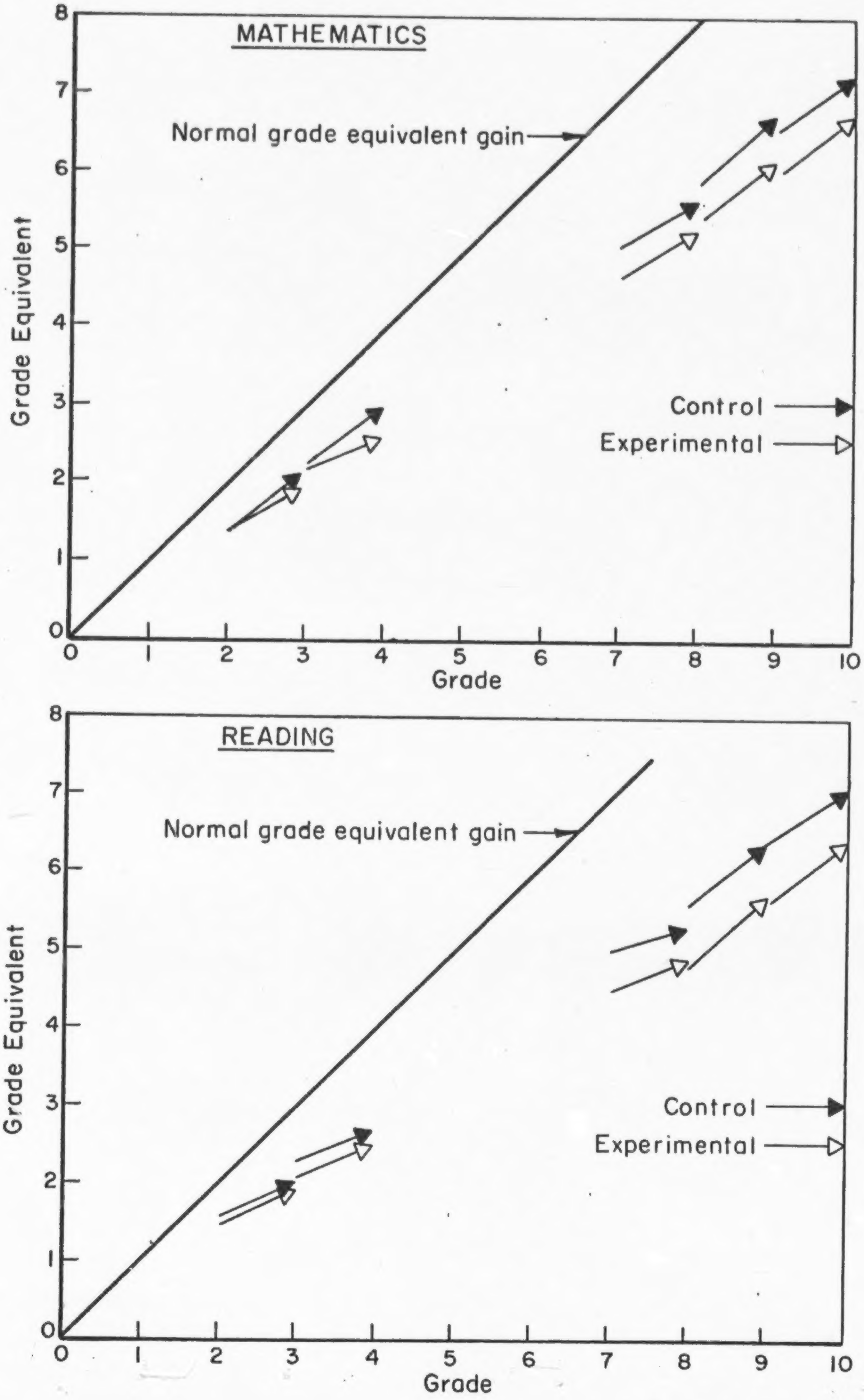


FIGURE 3. GRADE EQUIVALENT GAINS BY GRADE FOR MATHEMATICS AND READING

TABLE 31. OVERALL MEAN PRETEST AND POSTTEST VALUES (AND ASSOCIATED GRADE EQUIVALENTS) FOR EXPERIMENTAL (E) AND CONTROL (C) STUDENTS, BY SUBJECT AREA AND GRADE

MATHEMATICS

	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
	E	C	E	C	E	C	E	C	E	C	E	C
Pretest	68 (2)*	74 (3)*	28 (1.4)	31 (1.4)	44 (2.2)	51 (2.3)	43 (4.7)	49 (5.1)	39 (5.4)	45 (5.9)	46 (6.0)	53 (6.6)
Posttest	38 (1.2)	43 (1.5)	42 (1.9)	45 (2.1)	62 (2.6)	69 (3.0)	53 (5.3)	60 (5.7)	45 (6.2)	53 (6.8)	53 (6.8)	60 (7.3)
Raw Score Gain	-	-	14	14	18	18	10	11	6	8	7	7
GE Gain	-	-	0.5	0.7	0.4	0.7	0.6	0.6	0.8	0.9	0.8	0.7
Sample Size	1042	1082	1152	1106	1242	1115	1197	1089	1065	1088	1102	975

READING

	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
	E	C	E	C	E	C	E	C	E	C	E	C
Pretest	68 (2)*	74 (3)*	32 (1.5)	35 (1.6)	33 (2.1)	42 (2.3)	40 (4.5)	46 (5.0)	32 (4.8)	38 (5.6)	38 (5.6)	45 (6.4)
Posttest	57 (0.8)	62 (1.2)	51 (1.9)	55 (2.0)	49 (2.5)	55 (2.7)	45 (4.9)	50 (5.3)	39 (5.7)	45 (6.4)	45 (6.4)	51 (7.1)
Raw Score Gain	-	-	19	20	16	13	5	4	7	7	7	6
GE Gain	-	-	0.4	0.4	0.4	0.4	0.4	0.3	0.9	0.8	0.8	0.7
Sample Size	1047	1092	1242	1192	1317	1165	1211	1108	1151	1105	1093	996

* Stanine values

This is shown in the plot by the open-headed arrow having its initial and final points corresponding to the grade equivalent gain. The associated solid-headed arrow shows the grade equivalent gain for the control group. For third grade, mathematics, these pre and post grade equivalents for the control group are 2.3 and 3.0, respectively. In this case the experimental group gained, on the average, 0.4 grade equivalents, whereas the control group gained 0.7 grade equivalents. The standard gain in grade equivalents is equal to 1.0, and will occur whenever an arrow has a slope equal to that of the normal grade equivalent gain line shown in the plot. Slopes less than one indicate a relatively reduced rate of achievement; the smaller the slope the greater is the reduction in achievement rate.

An examination of the figure shows that the experimental groups started at lower pretest levels than did the control groups for grades 3, 7, 8, and 9, in both reading and mathematics. In no case are the arrows parallel to the standard progression line. More importantly, in no case is the slope of an arrow associated with the experimental group markedly greater than that of the arrow associated with the control group. In other words, the rate of achievement of the experimental groups is not improved to any noteworthy extent over that of the control groups.

Experimental Versus Control Groups: Second-Year Stability of Impact Results

Selection of Grade/Site/Subject Combination for Second-Year Testing

In order to assess the stability of the impact of the experimental program over the summer months, certain grade/site combinations were selected within which students from the experimental (E) and control (C) groups were given the reading and/or mathematics subtests of the evaluation test during the fall of the 1971-1972 school year.* The particular form of the evaluation test given was the same form as that given to these students during the original pretest administration in the fall of the 1970-1971 school year, with the exception of Grade 1 students. Grade 1 students received the same test (and form) as they received during the spring 1971 posttest administration.

The primary strategy for selecting grade/site/subject combinations for second-year testing was to make a judgement, on the basis of preliminary data analyses,** where a significant positive or negative difference between the E and C groups occurred on the basis of first year test results. This judgement was made by examining the differences between E and C mean raw score gains, relative to their estimated variance. Second year testing, then, was planned for those grade/site/subject combinations where the preliminary analyses indicated either a positive or negative impact. Both experimental and control students were to be tested in the selected grades, sites, and subjects. Finally, the students to be tested during the second year had to be full year students during the first year, with both pre and posttest scores for the first year.

---

* Specifically, five sites (Selmer, Las Vegas, Athens, Portland, Fresno) tested during the week of November 29, 1971 and four sites (Dallas, Anchorage, Grand Rapids, and Jacksonville) tested during the week of December 6, 1971.

** At the time when decisions had to be made concerning where second year testing was to occur, formal statistical regression analyses had not yet been accomplished.

In accordance with the above procedures, the grade/site/subject combinations for which second year testing occurred are identified in Table 32. Whereas there is some balance between subject area, most of the testing occurred in the Elementary grades. This result is consistent with the higher density of significant Pre-Post Model differences in favor of the E or C group in the Elementary grades than in the Junior High grades in the first year analyses.

TABLE 32. SITE/GRADE/SUBJECT COMBINATIONS WHERE SECOND-YEAR TESTING WAS ACCOMPLISHED

Site	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
	R	M	R	M	R	M	R	M	R	M	R	M
Selmer	X		X	X	X	X						
Athens	X	X			X	X	X	X				X
Dallas	X	X		X	X	X	X		X	X		
Anchorage	X	X	X	X	X				X			
Las Vegas			X	X	X	X						
Fresno		X		X		X						
Grand Rapids								X	X	X	X	
Portland		X		X		X						
Jacksonville	X	X	X	X	X	X						

R = Reading
M = Mathematics
X = Occurrence of second year testing

Stability of Impact Results

Table 33 shows the estimated means for which the difference was statistically significant for the experimental and control groups using the pre-post regression model to analyze the fall 1971 second-year test scores against the fall 1970 pretest scores. The table shows, for example, that in Grade 2, mathematics, at Dallas, the mean pretest raw score is equal to 23. This value represents the mean raw pretest score for those students of the experimental and control groups who took both the fall 1970 pretest and the second year test. The grade equivalent value that corresponds to this mean is shown in parentheses to be equal to 1.2. The estimated second-year test mean for the experimental group at the mean pretest value of 23 is shown to be 44 with a grade equivalent of 2.0. The corresponding estimated mean value for the control group is shown to be 31 with a grade equivalent value of 1.4. The difference between these regression estimates of second-year test means is given in units of raw scores and grade equivalents in the column labeled E/C Difference. In the present example, the raw score difference is seen to be +13 and the grade equivalency is seen to be +.6.

Complete lists of all regression analysis results associated with the second year fall test are given in Appendix S, whether such results were statistically significant or not. Also, associated t values are presented there. In addition, Appendix P provides pre- and second-year test means (and associated grade equivalents) and standard deviations, for each of the experimental and control groups.

Table 34 illustrates the stability of statistical outcomes of the Pre-Post Model analysis of spring 1971 posttest performance to fall 1971 second-year test performance. In Grade 1 Reading, all outcomes remained unchanged. In Grade 1 Mathematics, one nonsignificant group difference from the spring analysis became significant in favor of the C group in the fall analysis; one E outcome become a nonsignificant outcome; two C's became N's; and only two outcomes, an N at Anchorage and an E at Jacksonville, remained the same.

TABLE 33. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) SECOND-YEAR TEST MEANS FOR SITES SHOWING A SIGNIFICANT GROUP DIFFERENCE

Grade Subject	Site	Mean Pretest Value*	Estimated Second Year Test Means E	C	E/C Difference
Grade 1 Reading	Selmer	(3)	81(2.2)	73(1.8)	+8(+.4)
	Dallas	(1)	63(1.2)	73(1.8)	-10(-.6)
	Jacksonville	(2)	66(1.5)	57(0.8)	+9(+.7)
Grade 1 Math	Athens	(2)	47(1.7)	53(1.9)	-6(-.2)
	Jacksonville	(2)	46(1.7)	34(1.0)	+12(+.3)
Grade 2 Reading	Jacksonville	28(1.4)	57(2.1)	51(2.0)	+6(+.1)
Grade 2 Math	Dallas	23(1.2)	44(2.0)	31(1.4)	+13(+.6)
	Las Vegas	21(1.2)	37(1.6)	45(2.1)	-8(-.5)
	Jacksonville	22(1.2)	46(2.2)	40(1.8)	+6(+.4)
	Portland	40(1.8)	54(2.7)	56(3.1)	-2(-.4)
	Fresno	32(1.5)	49(2.3)	53(2.6)	-4(-.3)
Grade 3 Reading	Selmer	45(2.4)	72(3.4)	66(3.1)	+6(+.3)
	Dallas	27(1.8)	40(2.3)	34(2.2)	+6(+.1)
	Las Vegas	29(1.9)	47(2.4)	56(2.7)	-9(-.3)
Grade 3 Math	Selmer	58(2.4)	90(3.9)	79(3.3)	+11(+.6)
	Dallas	36(1.7)	57(2.4)	49(2.3)	+8(+.1)
	Athens	51(2.3)	78(3.2)	72(3.0)	+6(+.2)
	Jacksonville	37(1.7)	63(2.6)	55(2.4)	+8(+.2)
Grade 7 Math	Athens	41(4.6)	54(5.3)	47(4.9)	+7(+.4)
Grade 8 Reading	Grand Rapids	32(4.8)	43(6.2)	37(5.5)	+6(+.7)

* For Grade 1 the mean pretest value is given in stanines; all other raw score means in the table have their associated grade equivalents in parentheses.

TABLE 34. STATISTICAL OUTCOMES OF THE PRE-POST MODEL ANALYSIS OF SPRING (1971) AND FALL (1971) TEST PERFORMANCE FOR ALL GRADE/SITE/SUBJECTS FOR WHICH THERE WAS SECOND-YEAR TESTING

	Grade 1				Grade 2				Grade 3				Grade 7				Grade 8				Grade 9			
	Reading		Math		Reading		Math		Reading		Math		Reading		Math		Reading		Math		Reading		Math	
Site	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall
Selmer	E	E		E	E	N	N	E	E	E	E													
Athens	N	N	N	C					N	N	E	E	E	N	N	E							E	N
Dallas	N	N	E	N			E	E	E	E	E	E	N	N			E	N	N	N				
Anchorage	N	N	N	N	N	N	N	N	N	N							E	N						
Las Vegas					C	N	C	C	C	C	C	N												
Fresno			C	N			C	C			C	N												
Grand Rapids															C	N	E	E	C	N	N	N		
Portland			C	N			C	C			C	N												
Jacksonville	E	E	E	E	E	E	E	E	E	N	E	E												

* E denotes impact in favor of the E group
C denotes impact in favor of the C group
N denotes impact in favor of neither group

A brief summary of the number and kind of changes in Table 34 is presented in Table 35. This table shows a two-way classification of outcomes obtained for all of the 46 site/grade/subject combinations where second-year testing took place. The right-hand column of the table shows that 20 of the 46 combinations yielded a significant impact for the experimental groups as a result of the spring posttest analysis; 14 of the 46 combinations showed no impact as a result of the spring posttest analysis; and 12 of the combinations showed an impact in favor of the control group. The classification of impacts using the fall second-year test is shown by the bottom row of the table. This row shows 14, 27, and 5 impacts in favor of the experimental group, neither group, and the control group, respectively.

An examination of the main diagonal of the table shows that 13+12+4 = 29 of the 46 combinations yielded the same impact classification for both analyses; 17 combinations changed classifications between the spring and fall testing. Consequently, one measure of the stability of impact is given by the ratio 29/46, or 63 percent. This gives an estimate of the unconditional probability that a random selection from these 46 combinations will be classified the same way by the spring and fall analyses.

TABLE 35. CLASSIFICATION OF RESULTS TO SHOW STABILITY OF PROGRAM IMPACT

		Classification* Based on Fall 1971 Second-Year Test			
		E	N	C	Total
Classification* Based on Spring (1971) Posttest	E	13	7	0	20
	N	1	12	1	14
	C	0	8	4	12
	Total	14	27	5	46

* E denotes impact in favor of the experimental group
C denotes impact in favor of the control group
N denotes impact in favor of neither group

A more specialized examination of the E and C classifications can be made as follows. A total of 13 of the 20 combinations classified as E impacts by the spring posttests were also classified as E impacts by the fall tests. The ratio of 13/20 gives 65 percent as an estimate of the conditional probability that a randomly selected E-impact combination based on the spring posttest will also be classified as an E-impact combination by the fall second-year test. A total of only 4 of the 12 combinations classified as C impacts in the spring analyses are so classified in the fall. The ratio of 4/12 gives 33 percent as an estimate of the conditional probability that a randomly selected C impact from the spring analysis will also be classified as a C impact in the fall. It appears, then, that the outcomes in favor of the E group at the end of the 1970-1971 school year are more stable than the outcomes in favor of the C group. It also appears that the E impacts are more stable at the elementary than the secondary level, since only 4 of the 15 elementary E impacts changed from spring to fall, whereas 4 of the 5 secondary E impacts changed from spring to fall.

The nonsignificant group differences were the most stable of all. Twelve of the 14 N-impacts from the spring analyses remained the same; one became an E impact and one a C impact.

Because some of the students who took the first-year tests were not available to take the second-year tests, the results presented above are based on a sample of students that is "reduced" relative to the "original" sample of first-year students. In general, approximately 30 percent of the students in the original sample were lost. In order to determine whether the mean pretest values for the original and reduced samples were essentially equal, examinations of these means were made for both groups. These examinations indicated that over 90 percent of the 46 combinations that were posttested in the fall showed pretest means for the reduced sample within one raw score unit of the pretest mean of the original sample. Thus, it is concluded that students in the second-year analysis are equivalent to students in the first-year analysis, with respect to original entry level achievement, so that any change in impact from one analysis to the other cannot be explained on this basis.

Comparison Groups Analysis

In order to compare the performance of the experimental groups to the performance of students from within the same schools, the pool of replacement students at any grade/site was treated as a "comparison" (R) group for purposes of analysis. Comparison groups were posttested at any grade/site if there were judged to be at least 35 such students left at the end of the year. Table 36 identifies those grade/sites where comparison group posttesting occurred. Results of applying the Pre-Post Model to the analysis of experimental versus comparison group differences, and to control versus comparison group differences, are presented and discussed below. Also, experimental versus control group impacts are related to comparison versus control group impacts.

TABLE 36. SITES AND GRADES WHERE COMPARISON GROUPS WERE POSTTESTED

Site	Grade 1	Grade 2	Grade 3	Grade 7	Grade 8	Grade 9
Athens			X	X	X	X
Wichita	X	X	X	X	X	X
Rockland	X	X	X	X	X	X
Hartford				X	X	
McComb	X	X	X			
Jacksonville	X	X	X	X		X
Hammond	X	X	X			

X = Posttesting of comparison group

Control Versus Comparison Groups

Table 37 shows a summary of the significant differences between the control groups and the comparison groups. Grades where comparison group testing was not done are indicated by a dash in the table. The symbols R and C signify that the regression estimates of the posttest means differ significantly in favor of the comparison (R) or control groups (C), respectively. No significant difference is indicated by the symbol N. A complete listing of the regression results is given in Appendix T.

An examination of the table shows that comparison group testing was carried out for 58 grade/site/subject combinations. Among these combinations a total of 19 significant differences were found between the regression estimates of the posttest means for the control and comparison groups. Thus, significant differences are shown for approximately one-third of the testing combinations. Among these 19 differences, 14 (74 percent) favor the comparison group. The table also shows highly consistent results for a given site. For example, at Athens, Georgia, six of these groups gave significantly higher estimated posttest means for the comparison groups. In no case at Athens did the results favor a control group. The significant results at Rockland also favor the comparison groups, whereas the significant results consistently favor the control groups at Hammond and McComb. At Jacksonville, five results favor the comparison groups, and one result favors the control group. Although 12 comparison groups were tested at Wichita, no significant differences were found.

Experimental Versus Comparison Groups

Table 38 shows a summary of the significant differences between the experimental groups and the comparison groups. The symbols are the same as those shown in the preceding table except that the experimental group symbol E replaces the symbol C associated with control groups. These

TABLE 37. SUMMARY OF SIGNIFICANT DIFFERENCES BETWEEN CONTROL (C) GROUPS AND COMPARISON (R) GROUPS

	Grade											
	1		2		3		7		8		9	
Site*	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math
Athens	-	-	-	-	N	R	R	R	R	R	N	R
Hammond	C	N	N	N	C	N	-	-	-	-	-	-
Hartford	-	-	-	-	-	-	N	N	N	N	-	-
Jacksonville	R	R	N	R	R	R	N	C	-	-	N	N
McComb	N	N	N	N	C	C	-	-	-	-	-	-
Rockland	R	N	N	N	N	R	N	N	N	N	N	R
Wichita	N	N	N	N	N	N	N	N	N	N	N	N

* Grades where posttesting of comparison groups was not done are indicated by a dash; the symbol R (or C) signifies that the regression estimates of the posttest means favor the comparison group (or control) group; no significant difference is indicated by an N.

TABLE 38. SUMMARY OF SIGNIFICANT DIFFERENCES BETWEEN EXPERIMENTAL (E) AND COMPARISON (R) GROUPS

	Grade											
	1		2		3		7		8		9	
Site*	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math
Athens	-	-	-	-	N	R	N	N	R	R	N	R
Hammond	N	N	N	N	N	N	-	-	-	-	-	-
Hartford	-	-	-	-	-	-	N	N	N	N	-	-
Jacksonville	N	N	N	N	N	N	N	N	-	-	N	N
McComb	N	E	N	N	N	N	-	-	-	-	-	-
Rockland	R	R	N	R	N	N	R	R	R	N	N	N
Wichita	R	R	N	N	N	R	N	N	N	N	N	N

* Grades where posttesting of comparison groups was not done are indicated by a dash; the symbol R (or E) signifies that the regression estimates of the posttest means favor the comparison (or experimental) group; no significant difference is denoted by an N.

differences were obtained from the complete regression results presented in Appendix U. In addition, Appendix O provides pre and post comparison group means (and associated grade equivalents), along with a repeating of experimental group means, so that the interested reader can make comparisons between E and R groups in terms of pre and post means.

Among the 58 grade/site/subject combinations shown in the table, 13 favor the comparison groups, only one favors the experimental group, and the remaining 44 combinations favor neither group. The significant differences at Athens, Rockland, and Wichita all favor the comparison groups. The only significant difference in favor of the experimental group occurs at McComb in Grade 1 Mathematics.

Experimental Versus Control Group Impacts as Related to Comparison Versus Control Group Impacts

Table 39 shows ordered pairs of significant differences for the 58 site/grade/subject combinations where comparison groups were posttested. Consider the symbol ER associated with Athens, Grade 3 Mathematics. The first letter of this symbol indicates that the E group is significantly favored over the control group in the E versus C regression; the second letter indicates that the R group is significantly favored over the control group in the R versus C regressions. In general, the first symbol can be E, N, or C corresponding, respectively, to the experimental group, neither group, or the control group being favored by the experimental versus control regressions. Similarly, the second symbol can be R, N, or C corresponding, respectively, to the comparison group, neither group, or the control group being favored by the comparison versus control regressions. The first symbol is obtained from Table 38; the second symbol is obtained from Table 37.

An examination of this table shows that the experimental group is favored over the control group a total of 13 times (E occurs 13 times as the first symbol). The comparison group is also favored over the control group a total of 13 times (R occurs 13 times as the second symbol). Thus, the comparison groups are just as likely to be favored over the control groups as the treatment groups.

TABLE 39. PAIRED RESULTS OBTAINED FROM EXPERIMENTAL VERSUS CONTROL COMPARISONS AND COMPARISON VERSUS CONTROL COMPARISONS*

Site	Grade											
	1		2		3		7		8		9	
	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math	Read	Math
Athens	--	--	--	--	NN	ER	ER	NR	NR	NR	EN	ER
Hammond	CC	NN	NN	NN	CC	CC	--	--	--	--	--	--
Hartford	--	--	--	--	--	--	NN	NN	NN	NN	--	--
Jacksonville	ER	ER	EN	ER	ER	ER	NN	NC	--	--	NN	NN
McComb	NN	EN	CN	NN	NC	CC	CC	--	--	--	--	--
Rockland	NR	NN	NN	CN	NN	NR	CN	CN	CN	CN	NN	ER
Wichita	CN	NN	EN	NN	NN	CN	NN	NN	NN	NN	NN	NN

* The first symbol of each pair is E, N, or C corresponding, respectively, to the experimental group, neither group, or the control group being favored by the experimental versus control regressions; the second symbol of each pair is R, N, or C corresponding, respectively, to the comparison group, neither group, or the control group being favored by the comparison versus control regressions.

It is also interesting to note that in the 13 cases when the experimental group is favored over the control group, in 9 out of 13 cases (69 percent), the comparison group is also favored over the control group. This possibly suggests the operation of a "school factor" in certain cases of an apparent experimental impact.

Finally, control groups do better when paired against experimental groups than when paired against comparison groups. Thus, control groups are favored 14 times when paired against experimental groups, but control groups are favored only six times when paired against comparison groups.

Analysis of Special Treatment Groups

At two of the eighteen sites, Grand Rapids and Hartford, a number of additional remedial education programs were in progress during the 1970-71 school year for students in schools other than those designated as experimental and control schools. At Grand Rapids, students from Grades 1, 2, 3, 7, 8, and 9 were participating in four other programs: At Hartford, students from Grades 1, 2, and 3 were participating in two special treatment programs. Table 40 identifies these programs by site and grade. These programs are described breifly below. Following this description, the results of comparing the special treatment group with the control and experimental groups are presented and discussed.

Description of Special Treatment Groups at Grand Rapids

The Westinghouse Learning Centers (Grades 1-3) program was quite similar to WLC programs in the experiment. It departed from experimental treatments, however, in that the program was not in grades 7-9 and also the size of individual centers was smaller. In the WLC experimental programs, individual centers were designed to accommodate 100 children at a time; in the WLC "special treatment" groups, however, center capcity was 30-40 children. The composition of student enrollment was similar to that in the experimental schools.

The Combined Motivations and Education Systems (COMES) (Grades 7-9) instructional program was heavily machine orientated. The staff was about 80 percent paraprofessional and the overall staff-student ratio was approximately 1/12. One class per week was held in "achievement motivation", in which motivational values and incentives

TABLE 40. SPECIAL PROGRAMS AT GRAND RAPIDS AND HARTFORD

GRAND RAPIDS

Program	Grade 1	Grade 2	Grade 3	Grade 7	Grade 8	Grade 9
Reading Centers		X	X	X		
COMES				X	X	X
Westinghouse	X	X	X			
Project Read	X	X	X			

X = Grades where programs implemented.

HARTFORD

Program	Grade 1	Grade 2	Grade 3	Grade 7	Grade 8	Grade 9
Waverly School	X	X	X			
Project CONCERN	X	X	X			

X = Grades where programs were implemented.

were discussed. There were no other incentive or point system. Students were similar to experimental students in terms of underachievement. Racially, most COMES students were black, compared to a mixed population in the experimental program and predominantly white population in the control program.

Project Read (Grades 1-3 - Reading Only) is a program developed by Behavior Research Laboratories (BRL). It used Sullivan curricular materials. Regular school staffing was employed. This included one paraprofessional in each classroom, in addition to the teachers, resulting in a staff ratio of approximately 1/13. (Paraprofessionals were included in this manner in all "inner city" classrooms, consequently control classes were staffed in the same way.) No hardware or external incentive systems were used. Although students were in a special treatment group in reading only, they were tested in both subjects.

Reading Centers (Grades 2, 3, 7 - Reading Only) was not a program per se; rather it consisted of the use of special remedial reading teachers to teach underachieving students in small groups in their schools. This instruction, unlike the case of all other programs in the experiment, was supplementary to the school's normal reading programs. Each teacher was responsible for 2 or 3 schools. Scheduling of classes varied both as to times per week and size of the group. On the average, sessions were held with individual children 2-3 times per week.

Description of Special Treatment Groups at Hartford

The Waverly School (Grades 1-3) program consisted of a newly opened elementary school using the "open space" concept. It drew students from the same area as the experimental and control schools. Enrollment was governed by normal policies of the Hartford School District and was not related to the experiment. Student characteristics were similar to those in the experimental and control schools. In addition to the open space environment, individualized instruction and team teaching were used. No particular emphasis was placed upon hardware, other than might be expected in a new school and no external incentive systems were in use. Paraprofessionals supplemented certificated teaching staff. The overall staff-student ratio was approximately 1/11.

Project CONCERN (Grades 1-3) is a program to bus disadvantaged children from the same general area served by experimental and control schools to middle class, predominantly white schools in other sections of the city. These children received the normal instruction provided in most of the city's schools. The overall staff-student ratio was approximately 1/16. The program was started on an experimental basis and was considered to be operational during 1970-71.

Control Versus Special Treatment Groups at Grand Rapids

Table 41 shows the regression estimates of the posttest means for which the difference was significant for the control and special treatment groups at Grand Rapids, Michigan. The table shows, for example that in Grade 1 the special treatment (ST) program, Project Read, yields 69 for the regression estimate of the posttest mean corresponding to the pretest stanine value of 3. The regression line for the control group yields 60 for its estimated posttest mean. The difference between these two estimates is -9, in favor of the special treatment group, and is shown in the column labeled C/ST Difference.

An examination of the table shows that four out of five of the significant differences are in favor of the special treatment programs: Project Read, COMES, and the Reading Centers. Only in Grade 8-Mathematics does the control group perform significantly better than the special treatment, COMES. These five significant differences were found in a set of 24 comparisons between special treatment groups and control groups. Complete results for all comparisons are given in Appendix V.

Control Versus Special Treatment Groups At Hartford

Table 42 shows similar results for comparisons between special teatment groups and control groups at Hartford, Connecticut. The

TABLE 41. ESTIMATED POSTTEST MEANS FOR CONTROL (C) AND SPECIAL TREATMENT (ST) GROUPS FOR GROUPS SHOWING SIGNIFICANT DIFFERENCES AT GRAND RAPIDS, MICHIGAN

Subject	Grade	Special Treatment Program	Mean Pretest Value*	Estimated Posttest Means		C/ST Difference
				C	ST	
Reading	1	Project Read	(3)	60(1.0)	69(1.6)	-9(-.6)
	8	COMES	31(4.6)	34(5.0)	39(5.7)	-5(-.7)
Math	3	Reading Centers	43(2.1)	62(2.6)	68(2.9)	-6(-.3)
		Project Read	40(1.9)	59(2.5)	67(2.8)	-8(-.3)
	8	COMES	36(5.1)	46(6.2)	37(5.4)	+9(+.8)

* Grade equivalents are presented in parentheses except in Grade 1 where stanines are given.

TABLE 42. ESTIMATED POSTTEST MEANS FOR CONTROL (C) AND SPECIAL TREATMENT (ST) GROUPS FOR GROUPS SHOWING SIGNIFICANT DIFFERENCES AT HARTFORD, CONNECTICUT

Subject	Grade	Special Treatment Program	Mean Pretest Value*	Estimated Posttest Means*		C/ST Difference
				C	ST	
Reading	1	Waverly School	(2)	58(.9)	65(1.4)	-7(-.5)
Math	1	Project CONCERN	(2)	41(1.4)	33(.9)	+8(+.5)
	2	Waverly School	26(1.3)	42(1.9)	48(2.3)	-6(-.4)
	3	Waverly School	42(2.0)	62(2.6)	70(3.0)	-8(-.4)

* Grade equivalents are presented in parentheses except for the Grade 1 pretest where stanines are given.

special treatment program at Waverly school shows significantly better performance than the control groups for Grade 1-Reading and for Mathematics in Grades 2 and 3. The reverse result is shown for Project CONCERN where the control group is estimated to yield a higher achievement level for Grade 1-Mathematics. The four significant differences shown in this table were found in a set of 12 such comparisons at Hartford, Connecticut. Complete results are given in Appendix W.

Experimental Versus Special Treatment Groups At Grand Rapids

Table 43 shows significant results obtained from the regression lines fitted to the data for the experimental groups and the special treatment groups at Grand Rapids, Michigan. The regression lines are evaluated at the mean pretest value to obtain the estimated posttest means.

An examination of the table shows that in every case in which a significant difference occurs, the difference is in favor of the associated special-treatment programs: Project Read, Reading Centers, Westinghouse, or COMES. The ten significant differences are equally divided between reading and mathematics, and occur for Grades 1, 2, 3, and 7. Complete results on a set of 24 such comparisons at Grand Rapids are given in Appendix V.

Experimental Versus Special Treatment Groups At Hartford

Table 44 shows results similar to the preceding table for special treatment groups, Waverly School and Project CONCERN, at Hartford, Connecticut. The table shows that the regression estimates favor the special treatment groups over the experimental groups in every case in which a significant difference is judged to occur. Higher achievement is shown for three cases in reading, three cases in mathematics. Complete results for an additional set of six comparisons are given in Appendix W.*

---

* In addition to the regression results given in Appendices V and W for experimental vs. special treatment groups, Appendix Q provides a complete listing of pre and post means (and associated grade equivalents) and standard deviations.

TABLE 43. ESTIMATED POSTTEST MEANS FOR EXPERIMENTAL (E) AND SPECIAL TREATMENT (ST) GROUPS FOR GROUPS SHOWING SIGNIFICANT DIFFERENCES AT GRAND RAPIDS, MICHIGAN

Subject	Grade	Special Treatment Program	Mean Pretest Value*	Estimated Posttest Means*		E/ST Difference
				E	ST	
Reading	1	Project Read	(2)	52(<.6)	59(1.0)	-7(<-.4)
	2	Reading Centers	26(1.3)	46(1.8)	58(2.0)	-12(-.2)
	3	Reading Centers	31(2.0)	43(2.4)	52(2.6)	-9(-.2)
		Westinghouse	34(2.2)	45(2.4)	51(2.6)	-6(-.2)
	7	COMES	33(3.9)	38(4.4)	42(4.8)	-4(-.4)
Math	1	Project Read	(2)	30(.7)	37(1.2)	-7(-.5)
	2	Reading Centers	32(1.5)	44(2.0)	49(2.3)	-5(-.3)
	3	Reading Centers	40(1.9)	57(2.4)	65(2.7)	-8(-.3)
		Project Read	38(1.8)	56(2.4)	67(2.8)	-11(-.4)
	7	COMES	33(4.1)	41(4.6)	45(4.8)	-4(-.2)

* Grade equivalents are presented in parentheses except for the Grade 1 pretest where stanines are given.

TABLE 44. ESTIMATED POSTTEST MEANS FOR EXPERIMENTAL (E) AND SPECIAL TREATMENT (ST) GROUPS FOR GROUPS SHOWING SIGNIFICANT DIFFERENCES AT HARTFORD, CONNECTICUT

Subject	Grade	Special Treatment Program	Mean Pretest Value*	Estimated Posttest Means*		E/ST Difference
				E	ST	
Reading	1	Waverly School	(2)	54(.7)	65(1.4)	-11(-.7)
	2	Waverly School	26(1.3)	42(1.7)	52(1.9)	-10(-.2)
		Project CONCERN	29(1.4)	43(1.7)	53(2.0)	-10(-.3)
Math	1	Waverly School	(2)	33(.9)	44(1.6)	-11(-.7)
	2	Waverly School	22(1.2)	36(1.6)	46(2.2)	-10(-.6)
	3	Waverly School	40(1.9)	48(2.3)	69(3.0)	-21(-.7)

* Grade equivalents are presented in parentheses except for the Grade 1 pretest where stanines are given.

Summary of Results for the Special Treatment Groups

Table 45 provides a summary of the regression analysis results for the special treatment groups at Grand Rapids and Hartford. For any grade and subject area (reading or mathematics), the first column (labeled "E/ST") shows the results of comparing the experimental group with the various special treatment groups. Three symbols can appear: "E" for a statistically significant impact in favor of the experimental group, "ST" for a statistically significant impact in favor of the special treatment group, and "N" for no statistically significant difference between experimental and special treatment groups.

The second column in the table (labeled "C/ST"), for a given grade and subject area, shows the results of comparing the special treatment and control groups. Again, three symbols can appear, "C" indicating a statistically significant impact in favor of the control group, "ST" for a significant impact in favor of the special treatment group, and "N" denoting no significant differences between special treatment and control groups.

A dash in a given cell of the table indicates that the special treatment program was not implemented at that grade/subject combination.

For two of the programs, Westinghouse and Project CONCERN, the general outcome is no significant difference, with only one or two significant impacts when compared against either the experimental or control groups. The COMES program exhibits this same pattern, but to a lesser degree. Comparing COMES to the control groups in Grades 7, 8, and 9, 4 of the 6 comparisons yield no significant difference; and as compared to the experimental groups, again 4 of the 6 comparisons yield no significant difference.

On the other hand, the Waverly school program at Hartford shows promise of enhancing students' achievement. Of the 12 comparisons of this program with the experimental and control groups, in 8 of these 12 there was a statistically significant impact in favor of

TABLE 45. SUMMARY OF RESULTS FOR THE SPECIAL TREATMENT GROUPS*

A. GRAND RAPIDS

	Grade 1				Grade 2				Grade 3				Grade 7				Grade 8				Grade 9			
	Reading		Math		Reading		Math		Reading		Math		Reading		Math		Reading		Math		Reading		Math	
Programs	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST
Reading** Centers	--	--	--	--	ST	N	ST	N	ST	N	ST	ST	N	N	N	N	--	--	--	--	--	--	--	--
COMES	--	--	--	--	--	--	--	--	--	--	--	--	ST	N	ST	N	N	ST	N	C	N	N	N	N
Westinghouse	N	N	N	N	N	N	N	N	ST	N	N	N	--	--	--	--	--	--	--	--	--	--	--	--
Project Read**	ST	ST	ST	N	N	N	N	N	N	N	ST	ST	--	--	--	--	--	--	--	--	--	--	--	--

B. HARTFORD

	Grade 1				Grade 2				Grade 3				Grade 7				Grade 8				Grade 9			
	Reading		Math		Reading		Math		Reading		Math		Reading		Math		Reading		Math		Reading		Math	
Programs	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST	E/ST	C/ST
Waverly School	ST	ST	ST	N	ST	N	ST	ST	N	N	ST	ST	--	--	--	--	--	--	--	--	--	--	--	--
Project CONCERN	N	N	N	C	ST	N	N	N	N	N	N	N	--	--	--	--	--	--	--	--	--	--	--	--

* E denotes an impact in favor of the E group; C an impact in favor of the C group; ST an impact in favor of the special treatment group; N an impact in favor of none of the groups.

** Reading only taught in the special treatment program.

the Waverly school program. In 5 of the 6 comparisons with the experimental group, Waverly was favored, with 1 of the comparisons yielding no significant difference; in the 6 comparisons with the control group, Waverly was favored in 3, with 3 yielding no significant difference. In all four cases of no significant difference, however, the estimated differences in posttest means were nonetheless in favor of the Waverly program (see Appendix W), but not reaching statistical significance,

Thus, there is evidence that the Waverly program enhances student achievement, when compared against either the experimental or control groups, but more evidence when compared against the experimental group. It is interesting to note that the Waverly program does not emphasize hardware, and that no external incentive systems were in use, as discussed earlier. Rather, the focus was on the "open space" concept, emphasizing individualized instruction and team teaching.

The Reading Center program, and Project Read, both at Grand Rapids, also exhibit some evidence of having a positive effect on student achievement. Although, in both of these programs, only 5 of the 12 comparisons for each program were statistically significant in favor of the special treatment, for the remaining 7 comparisons, the regression estimates generally favored the special treatment programs (see Appendix V). That is, in these cases, the estimated difference in posttest means generally favored the special treatment group, but this difference was not large enough to reach the specified level of significance. It is again interesting to note that in both of these programs, there was no emphasis on hardware, and no external incentives were used. It is also interesting to note that in both of these programs, remedial instruction was given in reading only, as mentioned earlier.

In terms of comparing overall the experimental programs with the special treatment groups, the most conspicuous result shown in Table 45 is the absence of any positive impact in favor of the experimental treatment, out of 24 such comparisons in Grand Rapids, and out of 12 such comparisons in Hartford. However, the special treatment group was favored in 10 of the 24 comparisons at Grand Rapids, and in 6 of the 12 comparisons at Hartford.

In terms of comparing overall the control groups with the special treatment groups, the special treatment groups did not fare as well as when compared with the experimental groups. Out of 24 special treatment-control group comparisons in Grand Rapids, 4 yielded a significant impact in favor of the special treatments, with only 1 impact in favor of the controls. Out of the 12 special treatment-control group comparisons at Hartford, 3 yielded a significant impact in favor of the special treatments, and again with only 1 significant impact in favor of the controls.

Analysis of Attendance Data

In lieu of any variables which measure student motivation and attitudes toward school and learning, data were gathered on regular school attendance for students in the experimental and control groups wherever possible for the school years, 1970-71 and 1969-70 (the latter being considered as a "pre" measure or "entry level"). Attendance can be construed as at least an indicator of a student's willingness and desire to learn. The grade/sites for which data were collected for a Pre-Post Model analysis of attendance are identified in Table 46. The results of these analyses are presented and discussed here. Note that the attendance variable is expressed in units of percent days absent.

TABLE 46. SITES AND GRADES WHERE ATTENDANCE ANALYSES WERE PERFORMED

Site	Grade 1	Grade 2	Grade 3	Grade 7	Grade 8	Grade 9
Selmer		X	X	X	X	X
Athens		X	X	X	X	X
Wichita		X	X		X	X
Dallas		X	X	X	X	X
Rockland	X	X	X	X	X	X
Las Vegas				X	X	X
Fresno		X	X	X	X	X
McComb		X	X	X	X	X
Seattle		X	X	X	X	X

X = Attendance analysis performed

Table 47 shows the estimated 1970-71 percentage of school days absent for the experimental and control groups that showed significant differences between their rates of absenteeism. The attendance data were analyzed by regressing the percentage of days absent for each student during the 1970-71 school year against his percentage of days absent in 1969-70, as calculated from attendance records. The table shows, for example, that in the second grade at Dallas, the median* 1969-70 percentage of days absent for the combined experimental and control groups is equal to 10.25 percent. During the program, the estimated percentage obtained from the regression line fitted to the data for the E group is shown to be 1.29 percent. This value corresponds to the median value of 10.25 percent. The regression line, evaluated at 10.25 percent, yields an estimated mean of 7.72 percent for the C group. Thus, the absenteeism of the E group is estimated to be approximately 6.43 percentage points less than that of the C group, as shown in the column labeled E/C Difference. As shown in Table 47, a reduction in absenteeism of the E group over the C group occurred six times (Dallas, Grades 2, 7, and 8; Athens, Grades 8 and 9; and Selmer, Grade 9). These 6 cases may be compared with the treatment results shown in Table 27. Table 27 shows that the six grade/site combinations found to favor the E group on attendance yield four favorable and one unfavorable impact of the experimental treatment on posttest performance. Table 27 also shows that no favorable impacts and three unfavorable impacts on posttest performance were obtained for the four grade/site combinations shown in Table 47 that were found to favor the C group on attendance.

A complete listing of the attendance results is found in Appendix X. Among the 48 grade/site combinations found there, 24 show positive E/C differences and 24 show negative E/C differences.

---

* The median was used because the distribution of absenteeism was highly skewed.

TABLE 47. ESTIMATED MEAN 1970-71 PERCENTAGE OF SCHOOL DAYS ABSENT FOR EXPERIMENTAL (E) AND CONTROL (C) GROUPS SHOWING SIGNIFICANT DIFFERENCES

Grade	Site	Median 69-70 Percentage of Days Absent	Est. Mean 70-71 Percentage		E/C Difference
			E	C	
2	Dallas	10.25	1.29	7.72	-6.43
	Rockland	5.65	6.89	5.10	1.79
7	Dallas	8.70	7.14	10.62	-3.48
	Fresno	5.16	10.10	6.93	3.17
8	Dallas	5.97	5.83	14.12	-8.29
	Athens	6.17	5.50	8.93	-3.43
	McComb	3.97	7.35	4.76	2.68
9	Selmer	4.00	2.03	4.96	-2.93
	Athens	7.36	5.84	10.68	-4.84
	McComb	3.02	8.33	3.10	5.23

Thus, in general, it appears that the experimental treatment did not reduce absenteeism. However, wherever absenteeism was relatively low for the experimental groups vs. the control groups, the data suggest the possibility of enhanced experimental treatment impact on posttest performance.

Summary of Results and Conclusions

In this section a brief summary is given of the principal results presented in previous sections. Conclusions based upon these results are presented at the end of the section.

Summary of Principal Results

- Based upon 212 pre-post regression analyses of first year posttest performance in reading and mathematics of experimental and control students, 28 group differences (or 13 percent) were found in favor of the experimental group, 60 group differences (or 28 percent) were found in favor of the control group, and 124 group differences (or 59 percent) were not significant. In no grade, subject, or grade/subject combinations did the number of impacts favoring the experimental group exceed the number of impacts favoring the control group.
- Based upon 184 regression analyses involving extended regression models, again comparing the posttest performance in reading and mathematics of experimental and control students, 15 group differences (or 8 percent) were found in favor of the experimental group, 27 group differences (or 15 percent) were found in favor of the control group, and 142 group differences (or 77 percent) were not significant. Thus, the extended regression models yielded even a smaller percentage of impacts in favor of the experimental group than did the pre-post regression analysis.

- An overall comparison of the pre-post achievement gains of the experimental and control groups showed almost identical gains for the two groups, in reading and in mathematics, and at each grade level.
- Based upon 58 pre-post regression analyses of the experimental versus comparison group posttest performance in reading and mathematics, 13 significant differences were found in favor of the comparison group, and only one in favor of the experimental group.
- Based upon 46 pre-post regression analyses of second year test performance in reading and mathematics of experimental and control students, 13 (or 65 percent) of 20 first-year impacts in favor of the experimental program were maintained as measured by second-year test results.
- Of the 43 applications of the Pre-Post Model to attendance, comparing the experimental and control groups, 10 differences were significant. Six differences showed a lower rate of absenteeism for the experimental group than for the control group.
- Of the 24 pre-post regression analyses of experimental versus special treatment groups in reading and mathematics at Grand Rapids, 10 group differences were found to be significant. Of the 12 analyses of special treatment groups at Hartford, 6 group differences were found to be significant. All 16 significant differences were in favor of the special-treatment groups.

Conclusions

Several analyses of the data were performed, directed toward assessing the impact of performance incentive contracting on student achievement. Analyses were conducted at each site/grade/subject combination, as well as in the aggregate across several sites, and involved comparisons of several different groups with the experimental group.

As explicitly summarized above, all analyses point toward one outcome: the small amount of evidence in favor of the experimental groups. This outcome, however, must be evaluated in the light of three important points. First, the analysis evaluated performance incentive contracting as implemented by the six technology companies involved in the study. The evaluation makes no pretense of generalizing results and conslusions to the effectiveness of performance incentive contracting in general. That is, although efforts were made to select technology companies with widely varying educational approaches to implement their programs in a variety of locations, results and conclusions from this study are not to be generalized to the effectiveness of performance incentive contracting as implemented by other technology companies in other school districts. Indeed, the obtained outcomes for the six companies involved in the study could be different if they were to implement their program in a different way as a result of experiences gained during the experimental year.

Second, the analysis evaluated performance incentive contracting as implemented for a period of one year. The evaluation makes no pretense of generalizing results and conclusions to the effectiveness of performance incentive contracting implemented for more than one year.

Finally, the analysis evaluated performance incentive contracting using a standardized achievement test as the basis for assessing program impact. The evaluation makes no pretense of generalizing results and conclusions to the effectiveness of performance incentive contracting using some other method of assessing program impact.

Accordingly, the following conclusion is drawn:

- There is very little evidence that performance incentive contracting, as implemented by the technology companies at the 18 school districts in this study for a period of one year, had a beneficial effect on the reading and mathematics achievement of students participating in the experiment, as measured by a standardized achievement test.

RESULTS AND CONCLUSIONS: INCENTIVES ONLY SITES

In Battelle-Columbus' Interim Report on the Incentives Only sites (February 7, 1972)*, it was stated there that this interim report would constitute the final report on the Incentives Only sites, with the exception of reporting the results of additional analyses on data from these sites in this final report. It is the purpose of this section to review the results previously reported, and to present the results of additional analyses employing the Extended Variables Model, the Pre-Post Model with second year test scores in order to study stability of impact, and the Pre-Post Model applied to attendance data. For information on descriptions of the experimental programs at Stockton and Mesa, a description of the target population, reports on testing conditions, and all statistical appendices, the reader is referred to Battelle-Columbus' previous Interim Report.

Experimental Versus Control Groups: Regression Analyses of First-Year Results

The purpose of this section is to present the principal results based upon tests of significance of the differences between the experimental (E) and control (C) groups' estimated posttest means. As mentioned earlier, the analysis is confined to full-year students. Also, the analysis includes only those full-year students who have both a pretest and posttest score.

Pre-Post Model Results

Results are presented for reading and mathematics within each grade. Tables, and descriptive discussions of them, show the grade/site/subject combinations for which statistically significant differences were found, the estimated posttest means for each group, their differences, and the combined E and C pretest mean associated with the posttest differences. All analyses are done with raw scores; however, the grade equivalents of the raw scores are presented also to facilitate the interpretation of the results through the use of a commonly used metric.

---

* Op. cit., page 13

Table 48 shows the results obtained for Mesa and Stockton for those grade/subject combinations having a statistically significant group difference. As indicated by the table, for Grade 1, significant group differences occurred only in reading at Stockton. The difference found favored the experimental group. The table shows that a difference of +10 in the posttest means for Grade 1 reading is obtained when the fitted regression equations are evaluated at the combined mean of the pretest raw scores for the E and C groups. This combined mean is shown in Column 2 to have a stanine value of 4 and the individual regression estimates for the E and C groups are shown as 80 and 70 in Columns 3 and 4, respectively. The grade equivalents corresponding to 80 and 70 are shown in parentheses as (2.2) and (1.7), respectively, and the difference between these grade equivalents is shown as (+.5) in Column 5.

For Grade 2-Mathematics, the table shows that both Mesa and Stockton yielded significant differences between the posttest means when the regression lines were evaluated at the mean pretest value of 31. These posttest differences are seen to be +3 and +4 for Stockton and Mesa, respectively, and in both cases the differences favored the experimental group.

For Grade 2-Reading, no significant differences were found in either Mesa or Stockton. Likewise, for Grades 3, 7, and 8, no significant differences were found in Mesa or Stockton, in either reading or mathematics.

For Grade 9-Reading, Table 48 shows a significant negative impact for both Mesa and Stockton with a difference of -5 for the posttest means in each site as estimated by the regression equations. The grade equivalents show a gain from pretest to posttest of 1.2 for the control group at Stockton.

TABLE 48. ESTIMATED EXPERIMENTAL (E) AND CONTROL (C) POSTTEST MEANS AND DIFFERENCES BETWEEN THESE MEANS FOR THOSE GRADE/SITE/SUBJECT COMBINATIONS EXHIBITING A SIGNIFICANT GROUP DIFFERENCE

GRADE 1

	Site	Mean Pretest Value*	Estimated Posttest Means		E/C Difference
			E	C	
READING	Stockton	(4)	80 (2.2)	70 (1.7)	+10 (+.3)

* The stanine value corresponding to the combined E and C pretest raw score mean.

GRADE 2

	Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
			E	C	
MATH	Stockton	31 (1.4)	49 (2.3)	46 (2.2)	+3 (+.1)
	Mesa	31 (1.4)	43 (2.0)	39 (1.7)	+4 (+.3)

GRADE 9

	Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference
			E	C	
READING	Stockton	35 (5.2)	39 (5.7)	44 (6.4)	-5 (-.7)
	Mesa	46 (6.6)	48 (6.7)	53 (7.3)	-5 (-.6)

Summary of Pre-Post Model Results

Table 49 shows a concise summary of the principal results based upon the Pre-Post Model for Mesa and Stockton. The symbols, E and C, indicate the occurrence of a positive or negative difference, respectively, between the posttest means as estimated by the regression lines fitted to the pre-post data among those sites having a significant group difference. An empty cell indicates that no significant difference was found between the estimated posttest means for the E and C groups.

An examination of the table shows that Mesa had one positive impact (Grade 2-Mathematics) and one negative impact (Grade 9-Reading). Stockton had two positive impacts (Grade 1-Reading and Grade 2-Mathematics) and one negative impact (Grade 9-Reading). For Stockton and Mesa combined, the total number of positive impacts is 3; the number of negative impacts is 2; and the number of instances showing no significant differences is 19.

TABLE 49. SUMMARY OF PRE-POST MODEL RESULTS FOR EACH GRADE, AND SUBJECT AT MESA AND STOCKTON

ELEMENTARY

Site	Grade 1		Grade 2		Grade 3	
	Reading	Math	Reading	Math	Reading	Math
Mesa				E		
Stockton	E			E		

SECONDARY

Site	Grade 7		Grade 8		Grade 9	
	Reading	Math	Reading	Math	Reading	Math
Mesa					C	
Stockton					C	

Extended Variables Model Results

The Extended Variables Model was also applied in the analyses of the Incentives Only sites. Table 50 indicates the variables in this model for both sites.

TABLE 50. INDEPENDENT VARIABLES CONSTITUTING THE EXTENDED VARIABLES MODEL FOR EACH GRADE/SITE COMBINATION *

		Independent Variables					
Site	Grade	Group	Pretest	Student's Race	Father's Education	Total Family Income	Approve New Method
Stockton	All	X	X	X	X	X	X
Mesa	2	X	X	X	X		X
Mesa	8	X	X	X			
Mesa	1,3,7,&9	X	X	X	X	X	X

* Presence of variable in EVM is indicated by an "X".

Table 51 shows the results of changing from the Pre-Post Model to an Extended Variables Model for Stockton and Mesa. In the elementary grades, only the difference in favor of the E group in Grade 1-Reading at Stockton was maintained. In the junior high grades, the two differences in favor of the C group in Grade 9-Reading remained. Also, a difference in favor of the C group was obtained at Mesa with the EVM in Grade 7-Mathematics, where it had not been found with the PPM.

TABLE 51. COMPARISON OF SIGNIFICANT E/C GROUP DIFFERENCES FOUND USING THE PRE-POST MODEL VERSUS THE EXTENDED VARIABLES MODEL FOR EACH GRADE, SITE, AND SUBJECT

ELEMENTARY

	Grade 1				Grade 2				Grade 3			
	Reading		Math		Reading		Math		Reading		Math	
Site	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM
Stockton	E	E					E					
Mesa							E					

JUNIOR HIGH

	Grade 7				Grade 8				Grade 9			
	Reading		Math		Reading		Math		Reading		Math	
Site	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM	PPM	EVM
Stockton									C	C		
Mesa				C					C	C		

Stability of Impact Results

Second-year testing was carried out in Grades 2 and 3 at Mesa only.* There was one significant group difference found in these two grades based upon first year analyses and it was in favor of the experimental group in Grade 2-Mathematics. As Table 52 shows, this impact did not remain. No other impacts were found.

Analysis of Attendance Data

Application of the Pre-Post Model to attendance data was possible at Mesa in Grades 2, 3, 7, 8, and 9. As Table 53 shows, no significant differences between the experimental and control groups were found.

---

* Testing was accomplished during the week of November 29, 1971.

TABLE 52. REGRESSION RESULTS FOR SECOND YEAR-TESTING AT MESA

GRADE 2

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Reading	31 (1.5)	56 (2.0)	53 (2.0)	+3 (0)	1.019
Mathematics	32 (1.5)	44 (2.0)	44 (2.0)	0 (0)	-0.273

GRADE 3

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Reading	39 (2.3)	57 (2.8)	60 (2.9)	-3 (-.1)	-0.985
Mathematics	48 (2.3)	64 (2.7)	66 (2.8)	-2 (-.1)	-0.737

TABLE 53. REGRESSION RESULTS FOR ATTENDANCE AT MESA

Grade	Median 69-70 Attendance*	Estimated Mean 70-71 Attendance*		E/C Difference*	t Ratio
		E	C		
2	6.98	6.73	5.73	+1.00	0.993
3	7.74	7.39	7.15	+0.25	0.301
7	6.12	7.56	6.93	-0.63	0.582
8	6.27	5.95	7.86	-1.91	-1.286
9	7.32	7.73	9.53	-1.80	-1.494

* All entries are percentage days absent

Summary of Results and Conclusions

Summary of Results

- Based upon analyses in Grades 3, 7, 8, and 9, there is no evidence that the incentive programs at Stockton and Mesa had a beneficial effect on students' reading and mathematics standardized test achievement.
- Based upon analyses in Grade 2, there is little evidence that the incentive programs at Stockton and Mesa had a beneficial effect on students' mathematics standardized test achievement. The apparent impacts in Grade 2 mathematics at Stockton and Mesa in favor of the incentive programs were not found to be persistent under varying methods of analysis. For Grade 2 reading, there is no evidence that the incentive programs had a beneficial effect on students' reading standardized test achievement.
- Based upon analyses in Grade 1, there is some evidence that the incentive program at Stockton had a beneficial effect on students' reading standardized test achievement, but not at Mesa. For Grade 1, mathematics, there is no evidence at either Stockton or Mesa that the incentive programs had a beneficial effect on students' standardized test achievement.

Conclusion

Based on the above results the following conclusion is drawn:

- Overall, there is little or no evidence at Stockton and Mesa that the "Incentives Only" programs were beneficial to the students in reading or mathematics achievement, as measured by a standardized test.

APPENDIX A

TABLE A-1. NUMBER OF FULL-YEAR STUDENTS BY SITE, GROUP, AND GRADE

	Grade											
	1		2		3		7		8		9	
Site	Exp	Con	Exp	Con	Exp	Con	Exp	Con	Exp	Con	Exp	Con
Anchorage	50	62	53	54	56	57	51	71	51	69	58	53
Athens	78	82	78	94	87	80	85	79	81	75	77	76
Bronx	84	101	67	107	86	94	68	123	82	82	65	83
Dallas	87	86	89	89	88	89	83	75	89	84	86	88
Fresno	81	89	87	92	91	91	87	87	91	86	80	83
Gr. Rapids	82	78	91	80	92	77	92	85	87	89	84	71
Hammond	84	111	77	108	85	134	96	101	92	100	88	98
Hartford	70	93	69	93	74	76	79	70	95	90	93	126
Jackson	93	94	95	89	93	86	88	86	94	91	85	90
Las Vegas	69	47	66	51	82	34	83	54	84	52	71	55
McComb	72	47	71	44	71	53	72	55	75	47	73	66
Phila.	117	58	111	113	102	84	52	67	65	60	58	46
Portland	84	100	89	102	89	100	88	97	84	93	73	94
Rockland	101	80	91	79	112	82	102	73	100	68	87	75
Seattle	75	80	79	89	82	92	88	94	84	96	89	95
Selmer	94	38	100	46	103	47	94	54	88	52	98	40
Taft	64	83	67	99	75	95	65	79	51	85	47	51
Wichita	86	60	83	73	81	60	82	95	85	84	89	93

APPENDIX B

TABLE B-1 SURVEY RETURN RATE BY SITE, GROUP, AND GRADE

	Grade											
	1		2		3		7		8		9	
Site	Exp	Con	Exp	Con	Exp	Con	Exp	Con	Exp	Con	Exp	Con
Anchorage	43	52	41	58	49	61	34	56	31	65	40	51
Athens	68	63	72	74	68	68	48	69	60	59	58	63
Bronx	26	18	20	23	23	21	13	8	23	7	0	0
Dallas	68	44	71	42	69	31	50	47	38	*	41	63
Fresno	63	66	69	66	68	62	42	40	26	36	29	40
Gr. Rapids	59	45	46	42	52	50	52	49	35	50	42	34
Hammond	52	47	53	53	50	46	49	46	38	46	36	64
Hartford	14	31	26	37	16	42	26	30	22	37	37	30
Jackson.	82	68	75	75	70	69	84	67	75	70	72	70
Las Vegas	45	9	46	34	30	35	37	54	34	30	25	29
McComb	62	38	51	46	52	37	44	58	64	43	51	57
Phila.	23	0	22	0	27	0	9	17	15	33	4	19
Portland	51	52	52	59	47	56	46	64	44	49	29	68
Rockland	76	0	64	0	59	0	46	0	47	0	34	0
Seattle	48	60	45	60	50	56	40	42	39	31	35	40
Selmer	60	61	64	57	67	59	60	49	65	68	56	49
Taft	44	20	28	24	37	15	25	15	23	22	23	21
Wichita	56	45	64	50	70	40	72	78	54	80	58	57

* Dallas eighth grade produced a percentage return rate greater than one hundred. This can be only accounted for in terms of Dallas eighth grade adding more control students after the initial master list was created. The situation at Dallas was probably replicated at other sites; hence response rates should be interpreted cautiously.

APPENDIX C

PARENT QUESTIONNAIRE

Battelle Memorial Institute · COLUMBUS LABORATORIES

505 KING AVENUE COLUMBUS, OHIO 43201 · AREA CODE 614, TELEPHONE 299-3151 · CABLE ADDRESS: BATMIN

December, 1970

Dear Parent:

NATIONAL SCHOOL SURVEY

The purpose of these questions is to find out what parents think and feel about their children's schools and their children's progress in school.

We need your help in learning about how parents feel about their schools and their children's schoolwork so that changes can be made to improve education. Would you please take a few minutes to answer the questions in this survey? We need every parent's help. About 30,000 parents throughout the country are being asked these questions. After you have answered the questions, seal the form in the envelope and mail the questionnaire to Battelle Memorial Institute, Columbus Laboratories, in Columbus, Ohio.

Only the survey group in Columbus, Ohio, will see your answers to the questions. No one in your school district--teachers, principals, or anyone else--will see your answers. Your answers to the questions are confidential!

Thank you for your help.

Sincerely,

Dr. Kenneth W. Eckhardt

Dr. Kenneth W. Eckhardt
Educational Analysis
Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201

DIRECTIONS

Please circle the number of the statement which is closest to your answer.
Then write the number of your answer in the box at the right.

EXAMPLE

Should I circle the number of the answer I choose and write the number in the box?
Yes . . . ①
No . . . 2
[1]

THE FOLLOWING QUESTIONS ASK ABOUT YOUR FEELINGS TOWARD EDUCATION IN GENERAL

1. How good a job are the public schools doing in educating children?
A good job for all of the children 1
A good job for about half of the children 2
A good job for only a few of the children 3
Doesn't do a good job for any of the children 4
☐

2. How important is it for a child to get as much schooling as possible?
Very important . . . . 1 Not too important 3
Somewhat important . . . . . . 2 Not important at all 4
☐

3. Do you agree or disagree: Public school principals care what parents think?
Agree very much . . . . 1 Disagree somewhat 3
Agree somewhat . . . . 2 Disagree very much 4
☐

4. Do you agree or disagree: Public school teachers care what parents think?
Agree very much . . . . 1 Disagree somewhat 3
Agree somewhat . . . . 2 Disagree very much 4
☐

5. Do you feel parents have enough influence over what the schools do?
Yes . . . . 1 No. 2
☐

6. When children don't do well at school, which one of the following is most at fault?
Classes are too large 1
Teachers don't know how to teach. 2
School work is too hard 3
The children don't try 4
Parents don't help their children enough. 5
☐

7. Do you agree or disagree: School rules and regulations are too hard for parents to understand?
Agree very much . . . . 1 Disagree somewhat 3
Agree somewhat . . . . 2 Disagree very much 4
☐

THE FOLLOWING QUESTIONS ASK ABOUT YOUR FEELINGS TOWARD YOUR LOCAL SCHOOLS

8. Do you think that schools should be trying new teaching methods?
Yes . . . . 1 No. 2
☐

9. As far as you know, has your local school tried anything new in how children are taught?
Yes . . . . 1 No. 2
☐

10. Would you approve or disapprove of your children being taught by a new teaching method?
    Approve very much . . . . 1 Disapprove somewhat 3
    Approve somewhat . . . . 2 Disapprove very much. . . . . . 4 ☐

---

11. Is the child who brought this questionnaire home being taught by new teaching methods? ☐
    Yes . . . . 1 No. 2

---

12. From what you have seen or heard, do you think the new teaching method will be a success?
    My child is not being taught by any new method 1
    Yes, the new method will be very successful for my child 2
    Yes, the new method might be successful for my child . . . . 3 ☐
    No, the new method won't be successful for my child 4

---

13. Which kind of children do you think your child's teachers are most interested in teaching?
    Children who are fast learners. 1
    Children who are average learners 2
    Children who are slow learners 3
    All of the children. . . . . 4 ☐
    None of the children 5

---

14. How are your local schools doing in the following areas? **Circle a number for a, b, c, and d:**

	Excellent Job	Good Job	Fair Job	Poor Job	
a) School rules and regulations. . .	1	2	3	4	☐
b) School subjects taught	1	2	3	4	☐
c) Teacher's ability to teach	1	2	3	4	☐
d) Methods of teaching courses . .	1	2	3	4	☐

---

15. How satisfied are you with the local schools?
    Very satisfied . . . . 1 Dissatisfied . . . . 3 ☐
    Satisfied. . . . . 2 Very dissatisfied. 4

---

16. How often have you attended PTA or other meetings at your local school in the last year?
    Four or more times last year 1
    Two or three times last year. 2
    Once last year . . . . 3 ☐
    Never, last year 4

THE FOLLOWING QUESTIONS REFER TO THE CHILD WHO BROUGHT THIS QUESTIONNAIRE HOME FROM SCHOOL

17. Compared to most children how well is your child doing in school?
    Better than most children in the same grade. 1
    As good as most children in the same grade . . . . 2 ☐
    Not as good as most children in the same grade 3

---

18. Does your child get help in doing school work from someone in the family?
    Yes, often 1
    Yes, but only once in a while . . . . 2 ☐
    No, hardly ever 3

19. How often do you encourage your child to do well in school?

Very often	1
Somewhat often	2
Not too often	3
Hardly ever	4

☐

20. How much schooling **would you like to see** the child who brought this questionnaire home complete?

Some grade school (less than 8 years)	1
Complete grade school	2
Some high school	3
Complete high school	4
Some vocational or business school after high school	5
Complete vocational or business school after high school	6
Some college	7
Complete college	8
Professional or graduate school	9

☐

21. How much schooling do you think your child **will really get**?

Some grade school (less than 8 years)	1
Complete grade school	2
Some high school	3
Complete high school	4
Some vocational or business school after high school	5
Complete vocational or business school after high school	6
Some college	7
Complete college	8
Professional or graduate school	9

☐

THE FOLLOWING QUESTIONS ASK ABOUT YOUR FAMILY'S EDUCATION, JOBS, INCOME, AND OTHER FACTS

22. How much schooling have you and your husband or wife completed? Please **circle the answers.**

	Yourself	Husband or Wife
Some grade school (less than 8 years)	1	1
Completed grade school	2	2
Some high school	3	3
Completed high school	4	4
Some vocational or business school after high school	5	5
Completed vocational or business school after high school	6	6
Some college	7	7
Completed college	8	8
Graduate or professional school	9	9

23. Please circle **what grades all of your children** are enrolled in.

1st grade . . . 1	4th grade . . . 4	7th grade . . . 7	10th grade . . . 10	Technical
2nd grade . . 2	5th grade . . . 5	8th grade . . . 8	11th grade . . . 11	School . . . 13
3rd grade . . 3	6th grade . . . 6	9th grade . . . 9	12th grade . . . 12	College . . . 14

24. Have any of your children dropped out of school **before graduating** from high school?

Yes . . . . 1 No. . . . . 2 ☐

25. Do you have any children who have been in college?

Yes . . . . 1 No. . . . . 2 ☐

26. Is the husband or wife of the household currently employed?

	Husband	Wife
Yes, Full-Time	1	1
Yes, Part-Time	2	2
No	3	3
Not living in house or dead	4	4

27. Please briefly describe:

Husband's job (if any) ______________________________

Wife's job (if any)______________________________

---

28. From the list below, which category best describes husband's job, wife's job? Please circle the one closest answer for husband; then wife.

	Husband	Wife
Mainly unskilled work (examples: cafeteria or laundry work unskilled factory work, car washer)	1	1
Farmer's helper or farms other persons land	2	2
Mainly semiskilled work (factory machine operator, gas station worker, delivery man, foundry worker)	3	3
Skilled manual work (sheetmetal worker, appliance repairman, railroad switchman, welder, electrician)	4	4
Clerical or sales worker (retail store, department store, typist, post office, bank clerk)	5	5
Small business owner (car dealer, gas station, dry cleaning shop, small grocery store)	6	6
Professional (social worker, nurse, librarian, teacher)	7	7
Large business owner or manager (branch manager of grocery chain, business executive, etc.)	8	8
Professional (chemist, dentist, physician, lawyer)	9	9

---

29. How many of your children are living at home?____________ ☐

---

30. **Including yourself**, how many people **live** in your house and are what you call family?________________________ ☐

---

31. Please circle whether or not you have the following things in your home.

	Yes	No		Yes	No
Television set.	1	2	Telephone	1	2
Radio	1	2	Magazines	1	2
Daily newspaper	1	2	Dictionary	1	2
Washing machine	1	2	Encyclopedia	1	2
Clothes dryer	1	2	Children's books on Science and Nature	1	2

---

32. Please circle the total family income for last year. Include all sources of income (work, gifts, welfare, social security and so on).

Under $2000	1	Between $7000 and $7999	7
Between $2000 and $2999	2	Between $8000 and $8999	8
Between $3000 and $3999	3	Between $9000 and $9999	9
Between $4000 and $4999	4	Between $10,000 and $12,999 . . .	10
Between $5000 and $5999	5	Between $13,000 and $15,999 . . .	11
Between $6000 and $6999	6	Over $16,000	12

☐

---

33. If you had no money coming into the house, how long would your savings last?

Less than 1 week	1	1 month to 3 months	3
1 week to 1 month	2	3 months to 6 months.	4

☐

---

34. Have you or are you now receiving financial aid from county, state, or federal social welfare agencies?

Yes, I am now receiving aid	1
Yes, in the past but not now	2
No, never	3

☐

---

35. Please circle your family position.

Father	1	Foster parent	3	Other	5
Mother	2	Family relative	4		

☐

---

36. Please circle your family background.

American Indian . .	1	Eskimo-Aleut . . .	4	White or Caucasian	7
Black or Negro . . .	2	Oriental	5	Other (Please write in)	
Mexican-American .	3	Puerto Rican	6	____________	8

☐

---

37. Please circle your sex.

Male.	1	Female	2

☐

APPENDIX D

TABLE D-1. PERCENTAGE DISTRIBUTION OF FULL-YEAR STUDENTS' RACE, BY SITE AND GROUP.

Race	Anchor-age		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jackson-ville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
White	54	91	37	59	8	2	0	0	29	43	47	56	57	87	1	--	0	0
Black	18	0	63	41	42	46	100	98	11	3	41	37	41	12	86	--	100	100
Spanish Speaking	2	1	0	0	42	50	0	2	58	53	9	6	2	1	13	--	0	0
Other	26	8	0	0	8	2	0	0	2	1	3	1	0	0	0	--	0	0
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	--	100	100
Sample Size:	315	356	484	459	436	397	517	497	513	502	520	468	503	612	223	--	543	508
Response Rate:	99	97	99	94	96	67	99	97	99	95	98	97	96	94	46	--	99	95

Race	Las Vegas		McComb		Phila-delphia		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
White	44	47	6	49	0	--	98	98	100	--	61	88	88	92	1	--	40	52
Black	45	46	94	51	100	--	2	2	0	--	30	7	12	8	2	--	58	47
Spanish Speaking	9	5	0	0	0	--	0	0	0	--	0	0	0	0	97	--	2	1
Other	2	2	0	0	0	--	0	0	0	--	9	5	0	0	0	--	0	0
	100	100	100	100	100	--	100	100	100	100	100	100	100	100	100	--	100	100
Sample Size:	452	292	432	306	478	--	502	536	579	--	483	417	560	269	365	--	505	425
Response Rate:	99	100	99	98	95	--	99	91	99	--	97	76	97	97	99	--	100	91

6/7

APPENDIX E

TABLE E-1. MEDIAN TOTAL FAMILY INCOME BY SITE, GROUP, AND GRADE.

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids	
	E	C	E	C	E	C	E	C	E	C	E	C
1	8,250	15,437	6,055	7,000	6,500	5,250	2,336	2,427	5,857	6,857	6,833	7,571
2	8,000	16,342	6,714	5,700	7,500	5,250	3,285	2,125	7,142	6,687	5,416	7,142
3	10,500	15,250	5,333	6,375	7,187	6,450	2,638	2,431	6,562	6,750	5,000	7,250
7	16,727	16,950	3,642	4,625	----	----	3,750	2,236	5,000	6,250	7,642	8,357
8	12,583	15,843	3,214	4,625	----	----	3,400	2,687	6,875	6,125	4,333	8,000
9	15,166	13,812	3,857	4,165	----	----	3,900	2,843	5,666	7,625	8,250	8,750

	Hammond		Hartford		Jacksonville		Las Vegas		McComb		Philadelphia	
	E	C	E	C	E	C	E	C	E	C	E	C
1	8,800	8,875	----	----	6,714	3,437	----	----	2,843	6,750	3,375	----
2	9,750	8,375	----	----	5,714	2,875	----	----	2,062	5,666	3,500	----
3	8,333	8,500	----	----	6,285	5,833	----	----	2,375	9,500	4,500	----
7	9,250	9,300	----	----	2,375	3,400	----	----	3,333	2,583	----	5,500
8	9,000	8,625	----	----	3,291	5,250	----	----	1,613	2,791	----	4,750
9	8,250	10,727	----	----	3,500	4,642	----	----	2,196	2,285	----	3,000

	Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C
1	6,666	6,875	6,954	----	6,750	8,625	6,714	4,833	2,875	2,375	4,576	3,833
2	7,833	6,428	6,600	----	7,750	6,312	6,416	5,500	2,375	3,000	5,200	4,300
3	5,357	5,833	6,777	----	7,000	8,000	6,571	4,750	3,222	2,625	6,100	6,000
7	6,571	10,375	6,666	----	10,125	11,928	6,166	5,000	2,375	3,500	8,100	8,500
8	6,000	9,562	6,250	----	9,000	11,977	6,000	4,500	2,843	3,125	11,125	11,088
9	6,333	10,812	6,666	----	10,968	11,333	4,772	6,000	2,479	3,000	11,046	11,375

8/9

APPENDIX F

TABLE F-1. PERCENTAGE DISTRIBUTION OF FATHER'S EDUCATION, FOR FULL-YEAR STUDENTS, BY SITE AND GROUP

Educational Level	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jacksonville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
• Less than High School	35	23	64	51	58	62	51	62	64	55	55	40	48	57	63	62	41	52
• Completed High School	26	30	26	27	28	24	30	22	23	25	25	38	32	26	21	26	27	25
• More than High School	39	47	10	22	14	14	19	16	13	20	20	22	20	17	16	12	32	23
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Sample Size:	148	225	283	298	87	70	236	183	219	259	243	224	237	312	117	140	400	337
Response Rate:	46	61	58	61	19	12	45	36	42	49	46	47	45	48	24	29	73	63

Educational Level	Las Vegas		McComb		Philadelphia		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
• Less than High School	40	40	71	49	60	51	48	40	44	-	35	28	52	58	89	87	38	31
• Completed High School	35	40	15	27	27	28	31	38	39	-	32	31	33	30	4	7	30	31
• More than High School	25	20	14	24	13	21	21	22	17	-	33	41	15	12	7	6	32	38
	100	100	100	100	100	100	100	100	100	-	100	100	100	100	100	100	100	100
Sample Size:	165	65	175	124	99	39	239	328	326	-	222	268	339	156	96	73	286	231
Response Rate:	36	22	40	40	20	9	47	56	55	-	45	49	59	56	26	15	56	50

10/11

APPENDIX G

TABLE G-1. PERCENTAGE DISTRIBUTION OF FATHER'S OCCUPATION, FOR FULL-YEAR STUDENTS, BY SITE AND GROUP

Occupational Category	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jacksonville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
• Semi or Unskilled	31	14	42	41	48	49	65	73	48	46	62	56	53	46	67	76	55	64
• Skilled Manual	37	44	18	31	23	30	15	15	31	29	22	28	34	38	23	17	27	21
• White Collar Professional or Business	32	42	39	28	29	21	20	12	21	25	16	16	13	16	10	7	18	15
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Sample Size:	114	208	167	195	62	47	118	54	174	212	170	177	199	276	57	87	222	181
Response Rate:	36	57	34	40	14	8	23	10	34	40	32	37	38	42	12	16	40	34

Occupational Category	Las Vegas		McComb		Philadelphia		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
• Semi or Unskilled	68	47	35	40	64	57	41	30	39	--	35	34	45	60	58	63	33	25
• Skilled Manual	23	17	35	37	23	19	27	30	21	--	43	30	24	25	33	29	34	34
• White Collar Professional or Business	8	36	30	23	13	24	32	40	40	--	22	36	31	15	9	8	33	41
	100	100	100	100	100	100	100	100	100	--	100	100	100	100	100	100	100	100
Sample Size:	131	49	96	84	43	16	187	273	264	--	148	219	264	123	87	62	209	165
Response Rate:	29	17	22	27	8	25	37	46	44		30	40	46	44	23	13	41	35

12/13

APPENDIX H

ENTRY LEVEL ACHIEVEMENT DATA FOR FULL-YEAR STUDENTS

TABLE H-1. MEAN GEQ VALUES OF FULL-YEAR STUDENTS FOR READING, BY SITE, GROUP, AND GRADE

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jacksonville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
1*	3	4	2	3	-	2	1	1	4	3	2	3	3	3	2	2	2	2
2	1.5	1.8	1.5	1.5	1.3	1.5	1.4	1.2	1.6	1.5	1.3	1.6	1.5	1.6	1.3	1.5	1.4	1.3
3	2.2	2.4	2.2	2.4	1.9	2.1	1.9	1.7	2.3	2.4	2.1	2.2	2.2	2.4	1.7	2.3	2.0	1.9
7	5.6	5.6	4.2	4.6	3.1	3.0	3.4	3.7	3.8	4.5	3.8	4.6	4.9	5.8	4.4	4.8	3.8	4.2
8	5.7	7.1	3.9	4.4	3.8	3.7	3.3	3.9	4.3	5.0	4.2	5.2	5.0	6.4	4.9	4.4	4.3	4.6
9	7.1	8.0	4.5	5.5	-	-	4.8	4.5	4.4	5.0	5.0	5.7	6.6	7.8	4.5	5.2	4.9	4.9

	Las Vegas		McComb		Phila.		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
1*	2	3	1	3	2	2	3	4	4	5	3	4	2	3	2	2	3	3
2	1.3	1.3	1.3	1.6	1.3	1.3	1.7	1.9	1.7	2.0	1.5	1.8	1.8	1.5	1.5	1.7	1.5	1.9
3	1.9	2.0	1.8	2.4	1.8	1.5	2.2	2.4	2.2	2.5	2.0	2.7	2.4	2.4	2.2	2.3	2.1	2.4
7	5.2	4.8	4.3	4.1	3.1	3.2	4.9	5.7	5.2	6.7	5.3	6.6	6.0	5.2	3.8	4.6	4.8	4.6
8	5.8	5.5	3.9	4.6	3.5	4.6	5.0	6.0	5.6	8.0	5.6	6.4	6.7	7.1	4.2	4.8	5.3	6.2
9	6.6	7.3	4.6	5.8	3.7	3.8	6.4	6.9	6.9	9.8	6.4	7.1	6.4	7.4	5.3	5.5	6.4	7.3

---

* Stanine values are given for Grade 1.

15

TABLE H-2. MEAN GEQ VALUES OF FULL-YEAR STUDENTS FOR MATH BY SITE, GROUP, AND GRADE

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jackson-ville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
1*	3	4	2	3	-	2	1	1	4	3	2	3	3	3	2	2	2	2
2	1.4	1.5	1.4	1.4	1.2	1.2	1.2	1.2	1.4	1.5	1.5	1.5	1.4	1.5	1.2	1.3	1.2	1.2
3	2.2	2.4	2.2	2.4	2.2	1.9	1.6	1.6	2.2	2.4	2.0	2.2	2.2	2.4	2.0	2.3	1.8	1.7
7	4.9	5.6	4.2	4.8	4.5	4.4	4.1	3.7	4.2	4.5	4.3	4.6	5.1	5.6	4.6	4.7	4.2	4.5
8	5.6	6.2	4.5	4.6	4.4	4.4	3.7	4.3	4.5	5.4	4.9	5.4	5.4	6.4	5.6	5.1	5.1	5.4
9	6.6	6.7	5.2	5.9	-	-	5.2	5.4	4.8	5.3	5.6	5.9	6.6	7.5	5.3	6.0	5.7	5.7

	Las Vegas		McComb		Phila.		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
1*	2	3	1	3	2	2	3	4	4	5	3	4	2	3	2	2	3	3
2	1.1	1.2	1.2	1.3	1.3	1.2	1.7	1.7	1.5	2.0	1.4	1.5	1.5	1.6	1.3	1.4	1.3	1.5
3	1.3	1.4	1.6	2.4	1.7	1.7	2.4	2.6	2.3	2.7	1.9	2.6	2.4	2.6	2.3	2.3	2.0	1.9
7	4.9	4.7	4.5	4.4	3.9	3.5	4.9	5.6	5.3	6.4	4.9	5.9	5.6	4.9	4.5	5.3	4.6	4.4
8	5.4	5.7	5.1	5.9	4.1	5.4	5.9	6.4	6.0	7.8	5.9	6.4	6.8	7.5	4.8	5.7	4.9	5.7
9	6.4	6.7	5.7	6.8	4.4	4.6	6.6	6.8	6.7	9.2	6.7	6.6	6.8	7.6	6.4	5.7	5.9	6.6

---

* Stanine values are given for Grade 1.

TABLE H-3. PERCENTAGE OF FULL-YEAR STUDENTS BELOW GRADE LEVEL FOR READING, BY SITE, GROUP, AND GRADE

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jackson-ville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
2	100	76	96	88	98	98	87	100	82	92	99	88	92	81	100	85	94	99
3	100	86	97	89	100	98	100	100	96	85	100	89	91	77	100	89	97	99
7	88	87	99	95	100	100	100	100	100	99	100	95	99	71	96	90	100	99
8	91	69	100	100	100	100	100	100	100	96	100	95	98	70	91	95	100	97
9	81	61	100	90	100	100	99	98	99	98	100	100	90	66	100	94	97	95

	Las Vegas		McComb		Phila-delphia		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
2	100	100	98	90	100	97	80	67	82	43	90	70	65	93	98	86	93	64
3	100	100	100	88	100	100	90	85	100	60	99	57	87	83	100	96	100	88
7	97	96	100	98	100	100	100	91	93	54	91	59	71	86	100	98	100	98
8	93	89	100	100	100	93	100	91	89	45	90	81	69	73	100	100	90	90
9	86	72	100	92	100	100	89	90	76	36	89	76	88	69	98	87	90	71

TABLE H-4. PERCENTAGE OF FULL-YEAR STUDENTS BELOW GRADE LEVEL FOR MATHEMATICS, BY SITE, GROUP, AND GRADE

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jackson-ville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
2	94	75	87	83	97	99	94	100	89	86	80	89	87	78	100	94	100	97
3	100	85	96	86	100	100	100	100	99	90	98	94	92	78	100	91	99	99
7	98	86	100	97	98	100	100	100	100	100	100	99	96	80	98	96	100	99
8	96	94	100	100	100	100	100	100	100	95	98	99	99	83	96	98	99	100
9	96	92	100	95	100	100	100	100	99	98	97	99	99	83	100	98	99	99

	Las Vegas		McComb		Phila-delphia		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
2	96	98	96	98	97	100	63	58	79	42	91	76	77	67	98	87	90	72
3	100	100	100	94	100	100	92	74	87	62	99	72	75	67	99	80	99	100
7	100	100	99	98	100	100	100	89	93	61	96	74	63	88	100	94	99	99
8	97	98	97	96	100	95	97	91	93	48	99	90	70	67	100	93	100	99
9	98	93	100	88	100	100	99	98	89	39	92	94	91	70	100	93	99	90

APPENDIX I

PERCENTAGE DISTRIBUTION OF RESPONSES TO ITEM 10 (APPROVE NEW METHODS) OF PARENT QUESTIONNAIRE

19

TABLE I-1. PERCENTAGE DISTRIBUTION OF RESPONSES TO ITEM 10 (APPROVE NEW METHOD) OF PARENTS OF FULL-YEAR STUDENTS BY SITE AND GROUP

	Anchorage		Athens		Bronx		Dallas		Fresno		Grand Rapids		Hammond		Hartford		Jacksonville	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
Approve Very Much	58	43	56	54	78	63	75	80	65	54	62	61	62	47	60	69	78	67
Approve Somewhat	34	45	34	34	18	30	21	17	28	31	30	30	30	38	25	27	19	25
Disapprove Somewhat	5	9	7	8	2	6	2	2	6	10	6	6	4	11	8	2	2	3
Disapprove Very Much	3	3	3	4	2	1	2	1	1	5	2	3	4	4	7	2	1	5
	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Sample Size:	156	233	314	324	99	83	269	201	239	283	258	234	250	330	134	169	437	394
Response Rate:	49	64	65	67	22	14	52	39	46	54	49	49	48	51	28	31	78	74

	Las Vegas		McComb		Philadelphia		Portland		Rockland		Seattle		Selmer		Taft		Wichita	
	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C	E	C
Approve Very Much	67	56	68	48	68	55	57	44	57	--	63	45	64	76	50	71	58	45
Approve Somewhat	29	34	23	46	24	30	34	42	38	--	28	41	28	40	36	19	35	43
Disapprove Somewhat	4	4	4	2	5	2	7	9	3	--	4	9	4	7	8	6	3	7
Disapprove Very Much	0	6	5	4	3	13	3	5	2	--	5	5	4	7	6	4	4	5
	100	100	100	100	100	100	100	100	100	--	100	100	100	100	100	100	100	100
Sample Size:	177	68	206	151	103	40	244	339	333	--	231	286	356	161	121	97	295	239
Response Rate:	39	23	48	48	20	9	48	58	56	--	46	52	66	58	33	20	58	51

APPENDIX J

DESCRIPTION OF CERTIFICATION TESTS

Table J-1 lists all of the certification tests used for payment purposes. These tests are discussed, by grade, in this appendix.

TABLE J-1. LISTING OF CERTIFICATION TESTS *

Grades	Reading	Mathematics
1	• California Achievement Test 1970 Edition (Level 1) (A & B)	• California Achievement Test 1970 Edition (Level 1) (A & B)
2	• California Achievement Test 1970 Edition (Level 1) (A & B) • Metropolitan Achievement Test 1958 Edition (Primary I) (A & B) • Survey of Primary Reading Development (Forms A & B)	• California Achievement Test 1970 Edition (Level 1) (A & B) • Science Research Associates (Level 1-2) (C & D) • Stanford Achievement Test (Primary I) (W & X)
3	• California Achievement Test 1970 Edition (Level 2) (A & B) • Metropolitan Achievement Test 1958 Edition (Primary II) (A & B) • Survey of Primary Reading Development (Forms C & D)	• California Achievement Test 1970 Edition (Level 2) (A & B) • Science Research Associates (Level 2-4) (C & D) • Stanford Achievement Test (Primary II) (W & X)
7, 8, 9**	• California Achievement Test 1970 Edition (Level 4) (A & B) • Comprehensive Tests of Basic Skills (Level 3) (Q & R) • Iowa Tests of Basic Skills (1 & 2)	• California Achievement Test 1970 Edition (Level 4) (A & B) • Comprehensive Tests of Basic Skills (Level 3) (Q & R) • Iowa Tests of Basic Skills (1 & 2)

*Test levels and pre and posttest forms used are given in parentheses.

**The same three tests were used within 7th, 8th, and 9th grades.

21/22

Grade 1

In addition to being used as the Grade 1 evaluation posttest, the 1970 edition of the California Achievement Test, Level I, was considered the most satisfactory test for measuring the achievement of first grade students for payment purposes. The booklet is attractive and the printed pages and other materials are quite clear. The test has a good balance between readiness and achievement items related to vocabulary, comprehension, and number skills.

The normative information appeared sound and is based upon a stratification of school districts according to a geographic location, average enrollment per grade, and community type (determined by size and density of the community). The stratification originally yielded 84 cells but they were finally reduced to 53 because of the limited occurrence of only a few students per grade for some of the cells. Data are currently being analyzed, but at this time it appears that the groups are properly represented in the sample in proportion to their approximate ratio within the total population.

Although no reliabilities are reported for students at the beginning of first grade, the Kuder-Richardson Reliability Estimates (KR20) reported at mid-year are as follows:

	Reliability KR20	S.E.M.
Reading Total	.950	4.80
Vocabulary	.953	--
Comprehension	.759	--
Mathematics Total	.956	3.86
Computation	.947	--
Concepts and Problems	.904	--

These reliabilities reflect a problem with reading comprehension at this level, but in general the standard error of measurement is low compared to the total number of items (4.80 for 116 total reading items; 3.86 for 87 total mathematics items).

Finally, it is the one nationally standardized test which reports reading and mathematics grade equivalency scores below 1.0.

Although the first grade test in the California Achievement Test battery was judged to have a reasonably good balance between readiness and achievement items, there was reservation in using this test to measure the achievement of first graders at the beginning of the school year. It was expected, for example, that many of the students would be frustrated in attempting items which required the use of basic word knowledge and other reading skills which are not normally obtained until well into the mid-year of school. Accordingly, it was the recommendation of Battelle to depart from the existing testing schedule and delay the administration of this test until the mid-year for this group of children. This recommendation was not found to be acceptable, however, due to the constraint of the contractual agreements which existed between the Office of Economic Opportunity and the technology companies.

Because of these contractual agreements and in the absence of any other test which included a reasonable proportion of readiness items in addition to grade equivalencies reported below 1.0, it was decided to adhere to the schedule and conduct the administration at the beginning of the school year.

Grade 2

Level I of the California Achievement Test was also used as one of three certification measures for reading and mathematics at the second grade level. In addition to the advantages noted previously, this test was much more reliable at Grade 2 than Grade 1. For example, the Kuder-Richardson Reliability Estimates (KR20) reported for the test and the corresponding raw score standard error of measurements (S.E.M.) are as follows:

	Reliability KR20	S.E.M.
Reading Total	.968	3.50
Vocabulary	.964	2.84
Comprehension	.910	1.93
Mathematics Total	.953	3.13
Computation	.946	1.89
Concepts and Problems	.987	2.41

As reflected in Table J-1, this test was used to certify achievement in both reading and mathematics, for one third of the Grade 2 students. The other tests used for Grade 2 reading achievement were the Primary I level of the 1958 Edition of the Metropolitan Reading Achievement Test and the Survey of Primary Reading Development Test (Forms A and B). For measuring Grade 2 achievement in mathematics, the Scientific Research Associates Achievement Test, Level 1-2, and the Stanford Achievement Test, Primary I, were employed.

The 1958 edition of the Metropolitan Reading Achievement Test, Primary I, has long been a widely accepted measure of reading skills for beginning second grade students and it is considered to be one of the best measures of sight vocabulary (i.e., matching pictures with words and printed words with dictated words) and reading comprehension for this grade. This test yields three subscores: Word Knowledge, Word Discrimination, and Reading. The median corrected split-half reliability coefficients* and median raw score standard errors of measurement (S.E.M.) for each of these parts are as follows:

	Reliability Split-Half	S.E.M.
Word Knowledge	.90	2.3
Word Discrimination	.87	2.5
Reading (Comprehension)	.92	2.7

The reliabilities and the standard errors of measurement are acceptable considering that the total number of items for each part range between 35 and 45.

---

* The coefficients and standard errors of measurement reported are the median values obtained in a number of reliability studies.

One of the major strengths of this reading test is the careful analysis of the vocabulary used in materials available at the primary grades and the way this information was used to insure the validity of words used in the various sections of the test. Approximately 15,000 items were tested and analyzed before the final forms of the test were constructed.

The directions are clear; the format of the test is excellent; and the separately timed parts are sufficiently short to maintain the interest of the students throughout the administration.

There is relatively little information, however, about the norm sample other than a statement that the publishers attempted to represent the national school population with respect to: size of school system, geographic location, type of community (rural or urban), intelligence of pupils, and ethnic balance. Although nearly 500,000 students from 225 school districts in 49 states were tested, only approximately 25 percent of each class were included in the final norm group. There was no description of the procedures for selecting the original districts or reducing the sample. However, each school included was required to give the test to all pupils in regular classes in at least three consecutive grades.

The Survey of Primary Reading Development Tests were designed by Educational Testing Service as a "simple measure that would indicate rapidly and accurately the approximate level of a child's reading development". The test consists of six parts: Form Comparison, Word Form Comparison, Word Recognition, Sentence Recognition, Sentence Comprehension, and Story Comprehension. These parts reflect a broad scope of skills ranging from a readiness level to the higher skill levels not normally expected at the primary grades. The reliability of the test as estimated by the split-half method was found to be .91 and by the test-retest method, .88. In the test-retest method, the period of time separating the administrations was two weeks.

This test was judged to be outstanding for its content, scope, simplicity of directions, and overall format.

The major shortcoming of this test is the local sample used in the development of norms. Since the test was originally designed for students in the State of California, the entire norm sample consists of only students from that state. Thus, the out-of-state user must assume that the students in the norm population are representative of those in other parts of the country. This may be a more acceptable assumption than one might be led to believe because the sample did include: (1) students from a metropolitan city having low school mobility, high average socioeconomic status; (2) students from a residential urban community of high average socioeconomic status; (3) an urban area, half manufacturing and half residential with low mobility and low average measured intelligence; (4) an urban residential area with moderate mobility and about average measured intelligence; (5) an urban residential and business area of average socioeconomic and moderate mobility; and (6) a low average socioeconomic status community with low average measured intelligence and high mobility. Accordingly, the sample plan seems to reflect most types of urban or suburban communities in the country.

The Science Research Associates Achievement Test, Level 1-2, was used to measure the achievement of the second grade students in the areas of arithmetic concepts and arithmetic computations. The test of arithmetic concepts consists of 42 items. The reported Kuder-Richardson #20 reliability estimate for the test is .80 with a raw score standard error of measurement of 2.62. The test measuring arithmetic computations includes 52 items and has a much higher reliability, .95. The raw score standard error of measurement for this test is reported as 2.79.

It is not unusual to find a lower estimate of reliability in measuring arithmetic concepts because children at this age generally find it difficult to deal with the more abstract principles of arithmetic reasoning; thus, a greater variability in performance is evidenced. Computational skills, on the other hand, are generally concrete and, as a result, are handled much easier by the children. In view of this consideration, estimates of reliability and the reported raw score standard error of measurement for each were judged to be satisfactory.

The normative sample consisted of approximately 3,000 first grade students, 2,000 second grade students, and 1,000 third grade students from 250 schools. These schools were classified into one of eight geographic regions with each being designated as urban or rural. This norm design, as reported in the technical manual, however, does not yield much information concerning the appropriateness of the test for the population of students defined in this study.

The two subtests are well designed and have very clear directions and very short timed sections which should have the effect of increasing the reliability of the scores of students whose attention span is usually short. The directions for completing the items are clear and quite consistent. That is, the student is asked to mark an X on a numeral or write a numeral on a line. He is not confused by procedures calling for him to mark his answers in several different ways.

The Stanford Achievement Test, Primary II, in arithmetic was the fifth measure used at the second grade. This test also yielded two subscores, Arithmetic Computation and Arithmetic Concepts, and each consisted of 60 items and 46 items, respectively. Estimates of reliability were obtained through the use of the split-half technique. The estimate for the computation subtest was .93 with a raw score standard error of measurement of 2.09. The reliability estimate for the subtest measuring concepts was lower, .86, and it had a larger raw score standard error of measurement, 2.88.

The directions for completing the subtests are clear and easy to understand. The student is asked to either write his answer (in the computation test) or mark a cross in a circle (in the test on concepts).

There is relatively little information, however, about the norm sample other than a statement that the publishers attempted to represent the national school population with respect to: geographic location, type of school systems (there were eight types of schools which represented a combination of size and organizational structures) and the number of students desired per grade. Although the tests were given to about 850,000 students in 264 school districts drawn from the 50 states, only about 10,000 students per grade were actually included in the final normative sample. The procedure used for selecting the students in the sample and the weighting per grade was judged to be satisfactory.

In general, the Stanford Achievement Test in arithmetic is similar in content, ease of administration, and technical quality to the other arithmetic measures selected for this grade.

Grade 3

At Grade 3, Level 2 of the California Achievement Test was used in measuring achievement in both reading and mathematics. The other tests used for reading were the 1958 edition of the Metropolitan Reading Achievment Test, Primary II, and the Survey of Primary Reading Development, Forms C and D. For measuring achievement in mathematics, the Science Research Associates Achievement Test, Level 2-4, and the Stanford Achievement Test, Primary II, were employed.

Since these tests are higher level forms of those employed in the second grade, the discussion is limited to considerations of reliability. Table J-2 shows the reliability coefficients, raw score standard errors of measurement, and the number of items for each of the tests.

An examination of Table J-2 indicates the fairly high consistency of measurement that could reasonably be expected from the tests selected for measuring achievement in Grade 3.

TABLE J-2. RELIABILITY COEFFICIENTS, STANDARD ERRORS OF MEASUREMENT, AND NUMBER OF ITEMS FOR THE THIRD GRADE CERTIFICATION TESTS

Reading Tests	Reliability	S.E.M.	No. of Items
California Achievement Test Level 2 1970			
Reading Total	.957 (KR20)*	3.35	85
Vocabulary	.921 (KR20)	2.01	40
Comprehension	.933 (KR20)	2.61	45
Metropolitan Achievement Test, Primary II			
Reading Comprehension	.94 (Split-half)	2.8	51
Word Knowledge	.93 (Split-half)	2.2	37
Word Discrimination	.88 (Split-half)	2.3	35
Survey of Primary Reading Development, Forms C & D**	.86 (KR21)***	7.76	118
Mathematics Tests			
California Achievement Test Level 2 1970			
Arithmetic Total	.954 (KR20)	3.74	117
Computation	.939 (KR20)	2.80	72
Concepts and Problems	.900 (KR20)	2.40	45
Science Research Associates Test, Level 2-4			
Arithmetic Concepts	.82 (KR20)	2.46	39
Computations	.82 (KR20)	2.01	50
Stanford Achievement Test, Primary II			
Arithmetic Computation	.93 (Split-half)	2.57	60
Concepts	.90 (Split-half)	2.76	46

* Kuder-Richardson #20 formula for estimating reliability.

** This test contains Form Comparison, Word Recognition, Sentence Comprehension, Story Comprehension, and Pictorial Narrative.

***Kuder-Richardson #21 formula for estimating reliability.

Additional Supportive Information

An examination of the UCLA report, cited earlier, provided additional evidence of the quality of the tests selected at the primary level. In the report, existing standardized tests and test batteries were assessed in terms of the following criteria:

- How well the test measures the purported educational goal
- To what extent it is appropriate for students at that level
- To what degree the test can be easily utilized in the school
- To what extent the test is a sufficiently refined measurement tool.

The tests were rated with respect to these criteria on the basis of A - high, B - average, and C - low. The tests selected for this experiment received no rating below a B and a large number were rated A's in the category concerned with ease of administration. Since the tests are of professionally recognized national repute, the good ratings obtained in the UCLA study were not surprising, but rather were supportive of their repute. What was surprising, however, were the high ratings this study reported concerning the degree to which the tests could be easily utilized by school personnel.

Junior High School

At the junior high school level, three tests were used for measuring achievement in both reading and mathematics. These were the 1970 edition of the California Achievement Tests, Level 4, the Comprehensive Tests of Basic Skills, Level 3, and the Iowa Tests of Basic Skills. These same three tests were used at each of the grades 7, 8, and 9.

Level 4, California Achievement Test matches the excellence of the primary forms in terms of its test structure. It is well designed, contains clear directions, and the length of the separately timed subtests is judged appropriate for the population of students involved in the study. The technical quality of the test is also attested to by the high reliability estimates obtained; for example, on a population of seventh grade students. Table J-3 presents the reliability coefficients and standard errors of measurement for this group, as well as the number of items for each subtest. Similar data are also provided in the technical manual for Grades 6 and 8. The results are consistently similar to those reported below for Grade 7.

TABLE J-3. RELIABILITY COEFFICIENTS, STANDARD ERRORS OF MEASUREMENT, AND NUMBER OF ITEMS FOR A NATIONAL SAMPLE, CAT, LEVEL 4, GRADE 7

Test	Reliability	S.E.M.	No. of Items
Reading Total	.944 (KR20)	3.84	85
Reading Vocabulary	.911 (KR20)	2.54	40
Reading Comprehension	.889 (KR20)	2.85	45
Mathematics Total	.951 (KR20)	4.10	98
Mathematics Computation	.921 (KR20)	2.82	48
Mathematics Concepts and Problems	.898 (KR20)	2.97	50

The manual for this test battery has one of the best formats of any available. It is well laid out and easy for a relatively inexperienced examiner to follow. The answer sheets are designed well and should pose no problems for the students.

The Comprehensive Tests of Basic Skills, Level 3, published by the California Test Bureau, is a relatively recent addition to the test market and the content is quite current. Perhaps the greatest plus for this test battery is the characteristics of the normative population. The selected school districts were stratified according to size of school district enrollment, educational-economic index, and geographic region. Over 200,000 students participated in the norming of each form. But, the most significant factor for this study is the inclusion of students from each of the 23 largest districts in the country. Although no information is provided on the proportion of various ethnic minorities, it can be safely assumed that the nature of the stratification insures their participation in numbers proportionate to their representation in the total school population.

There is no question concerning the ease of administration as its structure is of comparable quality to that of the California Achievement Test.

The technical quality of the test is also attested to by the high reliability estimates obtained for the sixth grade students included in the national sample. Table J-4 presents the reliability estimates and standard errors of measurement for this group, as well as the number of items for each subtest. Although similar information is also available for seventh and eighth grades, the data are not significantly different from what is presented here.

TABLE J-4 RELIABILITY COEFFICIENTS, STANDARD ERRORS OF MEASUREMENT, AND NUMBER OF ITEMS FOR A NATIONAL SAMPLE, CTBS, LEVEL 3, GRADE 6

Test	Reliability	S.E.M.	No. of Items
Reading Total	.94 (KR20)	4.03	85
Reading Vocabulary	.91 (KR20)	2.70	40
Reading Comprehension	.89 (KR20)	2.97	45
Arithmetic Total	.95 (KR20)	4.33	98
Arithmetic Computation	.91 (KR20)	2.94	48
Arithmetic Concepts	.81 (KR20)	2.47	30
Arithmetic Applications	.85 (KR20)	1.91	20

The Iowa Test of Basic Skills has long been widely accepted as a highly refined measure of the goals and objectives of American education. It has been used in all types of school testing programs and at all levels with great success. It is also easy to administer, primarily because the tests for all grades are contained in one booklet.

It should be noted, however, that no separate measure of arithmetic computation is obtained for this test as there is for each of the other two. However, the problem solving subtest is very similar to the subtest measures of arithmetic computation in the other two tests, because nearly all of the problems involve the measurement of simple arithmetic operations. The difference in behavior measured is that the students do not respond to an operation sign, but to clue words which indicate what operations should be performed.

Table J-5 presents reliability data for Grade 8 only, but it is indicative of the quality of the instrument at other grade levels as well.

TABLE J-5. RELIABILITY COEFFICIENTS, STANDARD ERRORS OF MEASUREMENT, AND NUMBER OF ITEMS FOR THE ITBS, GRADE 8

Test	Reliability	S.E.M.	No. of Items
Vocabulary	.90 (Split-half)	3.0	48
Reading Comprehension	.93 (Split-half)	4.0	80
Arithmetic Concepts	.88 (Split-half)	2.2	48
Arithmetic Problem Solving	.79 (Split-half)	2.6	34
Arithmetic Total	.91 (Split-half)	2.1	82

In the normative sample approximately 20,000 students per grade were tested. The communities were first stratified by size of population, then by educational-economic index, and finally by geographic location. Although 15 of the 87 strata were not sampled because the schools were not willing to participate, the design was judged to be satisfactory.

APPENDIX K

MEANS, STANDARD DEVIATIONS, AND CORRELATIONS BETWEEN PRE- AND POSTTEST SCORES IN READING AND MATHEMATICS FOR FULL-YEAR STUDENTS WITH BOTH PRE- AND POSTTEST SCORES BY SITE AND GROUP

36

TABLE K-1. MEANS, STANDARD DEVIATIONS AND CORRELATIONS BETWEEN PRE- AND POSTTEST SCORES IN READING FOR FULL-YEAR STUDENTS WITH BOTH PRE- AND POSTTEST SCORES BY SITE AND GROUP

A. GRADE 1

Site		Experimental				Control			
		Mean	SD[(b)]	r[(c)]	N[(a)]	Mean	SD[(b)]	r[(c)]	N[(a)]
Anchorage	Pretest	74.46 (3)[(e)]	12.05	.436	46	91.12 (4)[(e)]	16.26	.404	52
	Posttest	58.33 (.9)	11.75			67.37 (1.5)	13.93		
Athens	Pretest	58.22 (2)	18.95	.464	64	71.21 (3)	21.60	.752	67
	Posttest	50.73 (.6)	11.69			57.55 (.9)	18.03		
Bronx[(d)]	Pretest	--	--	--	--	--	--	--	--
	Posttest	--	--			--	--		
Dallas	Pretest	44.25 (1)	18.42	.552	40	46.13 (1)	18.90	.592	46
	Posttest	52.12 (<.6)	14.29			49.04 (<.6)	14.98		
Fresno	Pretest	82.97 (4)	17.05	.297	60	81.95 (3)	17.39	.580	76
	Posttest	57.53 (.9)	11.94			56.28 (.8)	11.25		
Grand Rapids	Pretest	60.58 (2)	22.31	.604	60	74.46 (3)	19.67	.542	68
	Posttest	50.28 (<.6)	16.13			60.29 (1.0)	14.20		
Hammond	Pretest	74.69 (3)	18.16	.593	74	81.33 (3)	17.72	.624	97
	Posttest	60.32 (1.0)	17.03			74.42 (1.9)	15.04		
Hartford	Pretest	59.90 (2)	21.74	.647	40	61.28 (2)	20.15	.612	57
	Posttest	49.68 (<.6)	18.18			55.28 (.7)	14.18		
Jacksonville	Pretest	64.68 (2)	19.48	.777	90	56.79 (2)	17.63	.556	72
	Posttest	58.76 (1.0)	16.04			34.19 (<.6)	13.79		
Las Vegas	Pretest	62.20 (2)	19.53	.695	44	76.63 (3)	17.02	.423	41
	Posttest	48.16 (<.6)	11.92			60.10 (1.0)	14.40		
McComb	Pretest	50.09 (1)	17.89	.406	35	77.10 (3)	24.81	.788	47
	Posttest	56.09 (.8)	10.00			69.02 (1.6)	20.43		
Philadelphia	Pretest	62.99 (2)	18.18	.642	73	53.22 (2)	12.84	.289	51
	Posttest	52.63 (.6)	11.84			55.39 (.7)	13.20		
Portland	Pretest	79.06 (3)	20.41	.814	68	88.00 (4)	14.82	.419	64
	Posttest	58.90 (1.0)	18.18			72.72 (1.8)	11.82		
Rockland	Pretest	85.27 (4)	14.40	.662	82	93.53 (5)	15.35	.593	75
	Posttest	68.57 (1.6)	12.96			76.41 (2.0)	13.27		
Seattle	Pretest	76.29 (3)	12.19	.600	55	88.51 (4)	17.17	.720	65
	Posttest	58.42 (.9)	14.47			77.69 (2.1)	19.06		
Selmer	Pretest	70.00 (2)	17.46	.743	91	73.42 (3)	21.97	.770	33
	Posttest	67.07 (1.5)	14.38			56.42 (.8)	13.59		
Taft	Pretest	54.33 (2)	21.99	.528	51	61.92 (2)	17.70	.495	72
	Posttest	60.37 (1.0)	15.61			63.35	19.73		
Wichita	Pretest	73.24 (3)	12.29	.294	71	73.62 (3)	18.25	.728	56
	Posttest	53.34 (.6)	15.31			69.07 (1.6)	14.92		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

(d) At Bronx the E group did not take the evaluation pretest.

(e) Stanine values are given for Grade 1.

TABLE K-1. (Continued)

B. GRADE 2

Site		Experimental				Control			
		Mean	SD[(b)]	r[(c)]	N[(a)]	Mean	SD[(b)]	r[(c)]	N[(a)]
Anchorage	Pretest	31.31 (1.5)	8.41	.350	48	40.73 (1.8)	13.02	.488	44
	Posttest	60.46 (2.1)	13.40			61.11 (2.1)	15.51		
Athens	Pretest	32.15 (1.5)	10.17	.642	75	34.27 (1.5)	12.38	.724	85
	Posttest	50.71 (1.9)	15.68			52.05 (1.9)	18.52		
Bronx	Pretest	24.74 (1.3)	8.42	.409	42	31.06 (1.5)	10.90	.468	52
	Posttest	38.21 (1.6)	17.13			55.90 (2.0)	12.96		
Dallas	Pretest	29.74 (1.4)	16.22	.679	73	20.37 (1.2)	9.30	.468	63
	Posttest	39.86 (1.7)	16.83			30.30 (1.4)	11.75		
Fresno	Pretest	36.56 (1.6)	11.35	.613	81	33.83 (1.5)	13.01	.594	82
	Posttest	54.78 (2.0)	12.67			50.43 (1.9)	14.86		
Grand Rapids	Pretest	24.14 (1.3)	8.97	.575	71	35.04 (1.6)	13.20	.580	72
	Posttest	44.11 (1.8)	15.72			56.92 (2.0)	15.77		
Hammond	Pretest	31.04 (1.5)	11.44	.589	75	36.58 (1.6)	16.15	.744	96
	Posttest	50.63 (1.9)	16.26			57.04 (2.0)	17.47		
Hartford	Pretest	22.54 (1.3)	8.70	.368	54	30.55 (1.5)	16.29	.655	67
	Posttest	39.81 (1.7)	13.65			53.72 (2.0)	17.26		
Jacksonville	Pretest	29.35 (1.4)	11.89	.601	93	24.69 (1.3)	8.93	.354	77
	Posttest	51.84 (1.9)	17.78			42.45 (1.7)	15.85		
Las Vegas	Pretest	24.33 (1.3)	7.46	.626	49	21.71 (1.3)	6.87	.584	38
	Posttest	32.69 (1.5)	10.15			43.55 (1.8)	16.82		
McComb	Pretest	24.90 (1.3)	10.72	.568	43	35.16 (1.6)	10.59	.575	31
	Posttest	44.29 (1.8)	14.81			59.35 (2.1)	11.74		
Philadelphia	Pretest	25.15 (1.3)	8.07	.586	82	22.42 (1.3)	8.22	.497	65
	Posttest	41.66 (1.7)	14.84			40.85 (1.7)	12.21		
Portland	Pretest	39.05 (1.7)	14.64	.769	84	45.61 (1.9)	14.58	.538	80
	Posttest	56.10 (2.0)	15.31			66.34 (2.2)	12.06		
Rockland	Pretest	38.06 (1.7)	10.92	.585	80	54.25 (2.0)	16.17	.582	67
	Posttest	60.61 (2.1)	14.84			71.76 (1.8)	7.73		
Seattle	Pretest	34.36 (1.5)	11.20	.528	69	42.97 (1.8)	15.35	.515	71
	Posttest	54.61 (2.0)	15.57			66.20 (2.2)	11.35		
Selmer	Pretest	42.94 (1.8)	16.04	.635	92	34.41 (1.5)	11.11	.641	41
	Posttest	66.10 (2.2)	11.62			55.10 (2.0)	16.05		
Taft	Pretest	32.82 (1.5)	7.32	.569	61	40.10 (1.7)	11.91	.592	89
	Posttest	47.90 (1.9)	12.37			58.85 (2.1)	13.00		
Wichita	Pretest	32.18 (1.5)	10.48	.645	65	45.11 (1.9)	18.02	.667	72
	Posttest	57.48 (2.0)	13.71			61.90 (2.2)	15.27		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

TABLE K-1. (Continued)

C. GRADE 3

Site		Experimental				Control			
		Mean	SD[b]	r[c]	N[a]	Mean	SD[b]	r[c]	N[a]
Anchorage	Pretest	35.56 (2.2)	9.25	.425	50	48.24 (2.4)	18.31	.664	45
	Posttest	59.10 (2.8)	11.74			66.49 (3.1)	14.40		
Athens	Pretest	35.40 (2.2)	14.15	.634	85	42.70 (2.4)	17.86	.805	71
	Posttest	53.42 (2.6)	19.09			55.83 (2.7)	20.61		
Bronx	Pretest	29.25 (1.9)	10.52	.542	64	32.98 (2.1)	12.67	.579	55
	Posttest	46.08 (2.4)	15.63			54.76 (2.7)	16.35		
Dallas	Pretest	27.69 (1.9)	9.44	.443	71	24.24 (1.7)	8.77	.601	76
	Posttest	37.58 (2.2)	14.14			28.17 (1.9)	11.04		
Fresno	Pretest	39.94 (2.3)	13.70	.721	78	44.54 (2.4)	16.19	.725	85
	Posttest	51.97 (2.6)	15.86			56.88 (2.8)	16.77		
Grand Rapids	Pretest	33.12 (2.1)	12.73	.729	81	37.88 (2.2)	18.02	.794	67
	Posttest	44.72 (2.4)	17.27			50.12 (2.5)	17.26		
Hammond	Pretest	35.37 (2.2)	16.26	.784	74	45.77 (2.4)	20.42	.792	111
	Posttest	48.62 (2.5)	18.10			61.26 (2.9)	18.05		
Hartford	Pretest	24.96 (1.7)	9.59	.626	53	39.00 (2.3)	18.97	.745	46
	Posttest	44.89 (2.4)	15.06			54.96 (2.7)	15.80		
Jacksonville	Pretest	30.36 (2.0)	12.87	.714	91	29.38 (1.9)	11.29	.808	76
	Posttest	44.40 (2.4)	17.14			37.88 (2.2)	15.97		
Las Vegas	Pretest	27.74 (1.9)	8.83	.486	68	30.03 (2.0)	8.03	.447	31
	Posttest	38.41 (2.2)	11.73			47.94 (2.5)	15.82		
McComb	Pretest	25.54 (1.9)	7.93	.641	69	49.76 (2.4)	15.63	.768	51
	Posttest	40.12 (2.3)	15.04			67.12 (3.2)	14.89		
Philadelphia	Pretest	26.36 (1.8)	11.39	.757	66	20.32 (1.5)	4.91	.159	50
	Posttest	39.15 (2.3)	17.49			31.94 (2.1)	9.37		
Portland	Pretest	37.75 (2.2)	15.77	.683	79	44.69 (2.4)	18.89	.716	77
	Posttest	55.85 (2.7)	17.90			62.91 (3.0)	15.84		
Rockland	Pretest	34.99 (2.2)	9.63	.589	92	53.93 (2.5)	21.61	.793	75
	Posttest	58.89 (2.8)	16.11			69.35 (3.3)	14.40		
Seattle	Pretest	30.93 (2.0)	10.15	.350	72	60.05 (2.7)	17.41	.680	79
	Posttest	48.07 (2.5)	15.39			70.84 (3.4)	12.31		
Selmer	Pretest	44.32 (2.4)	16.86	.655	88	46.36 (2.4)	17.36	.742	45
	Posttest	67.15 (3.2)	12.38			58.58 (2.8)	17.31		
Taft	Pretest	38.32 (2.2)	11.26	.650	73	40.90 (2.3)	14.27	.718	83
	Posttest	52.81 (2.6)	13.43			54.05 (2.7)	15.51		
Wichita	Pretest	32.02 (2.1)	10.55	.610	63	44.17 (2.4)	16.67	.721	42
	Posttest	43.71 (2.4)	13.95			52.79 (2.6)	16.74		

[a] The number of students is indicated in the column labelled "N".

[b] SD = standard deviation

[c] r = pre-post linear correlation coefficient

TABLE K-1. (Continued)

D. Grade 7

Site		Experimental				Control			
		Mean	SD[(b)]	r[(c)]	N[(a)]	Mean	SD[(b)]	r[(c)]	N[(a)]
Anchorage	Pretest	52.67 (5.6)	13.20	.772	46	54.16 (5.6)	14.20	.648	57
	Posttest	58.54 (6.0)	12.78			59.37 (6.0)	15.31		
Athens	Pretest	34.73 (4.2)	13.59	.776	64	41.14 (4.6)	16.59	.818	70
	Posttest	39.98 (4.5)	18.28			41.86 (4.8)	16.57		
Bronx	Pretest	23.56 (3.1)	7.45	.560	44	22.62 (3.0)	6.89	.650	12
	Posttest	31.14 (3.8)	11.84			32.25 (3.8)	9.06		
Dallas	Pretest	26.79 (3.4)	9.91	.632	57	30.12 (3.7)	9.44	.690	43
	Posttest	27.91 (3.5)	11.11			28.95 (3.6)	11.99		
Fresno	Pretest	30.64 (3.8)	8.10	.624	83	40.45 (4.5)	11.71	.689	64
	Posttest	36.07 (4.3)	10.76			44.11 (4.9)	13.73		
Grand Rapids	Pretest	32.45 (3.8)	11.62	.704	59	40.59 (4.6)	15.45	.810	71
	Posttest	37.85 (4.4)	12.79			45.48 (4.9)	15.36		
Hammond	Pretest	44.04 (4.9)	12.37	.785	88	55.91 (5.8)	19.04	.913	96
	Posttest	51.34 (5.5)	13.70			62.22 (6.4)	17.73		
Hartford	Pretest	37.94 (4.4)	13.78	.831	71	41.93 (4.8)	17.72	.932	58
	Posttest	44.97 (4.9)	15.13			48.41 (5.2)	19.08		
Jacksonville	Pretest	31.03 (3.8)	11.54	.716	75	35.34 (4.2)	12.87	.808	71
	Posttest	33.65 (4.1)	14.30			39.89 (4.5)	15.15		
Las Vegas	Pretest	47.83 (5.2)	12.84	.787	64	43.02 (4.8)	13.95	.793	46
	Posttest	55.61 (5.8)	13.66			48.43 (5.2)	14.78		
McComb	Pretest	36.13 (4.3)	15.37	.900	69	33.73 (4.1)	14.74	.865	49
	Posttest	42.93 (4.8)	17.28			38.27 (4.4)	15.28		
Philadelphia	Pretest	23.58 (3.1)	7.86	.318	45	24.56 (3.2)	9.32	.529	39
	Posttest	27.56 (3.5)	8.15			29.79 (3.7)	13.65		
Portland	Pretest	44.69 (4.9)	10.57	.716	81	55.27 (5.7)	11.82	.825	97
	Posttest	51.30 (5.5)	12.93			59.92 (6.2)	14.31		
Rockland	Pretest	[illegible]8.21 (5.2)	16.05	.693	80	65.54 (6.7)	18.46	.883	59
	Posttest	[illegible].12 (5.[illegible])	15.58			72.14 (7.3)	16.87		
Seattle	Pretest	49.33 (5.3)	13.55	.786	70	64.06 (6.6)	16.31	.817	82
	Posttest	54.46 (5.6)	16.88			66.90 (6.7)	16.89		
Selmer	Pretest	58.14 (6.0)	17.16	.888	84	47.78 (5.2)	16.32	.851	49
	Posttest	64.69 (6.6)	16.65			53.27 (5.6)	17.46		
Taft	Pretest	31.85 (3.8)	10.74	.801	59	40.89 (4.6)	13.10	.858	61
	Posttest	35.80 (4.3)	10.88			44.90 (4.9)	13.93		
Wichita	Pretest	42.57 (4.8)	13.80	.796	72	41.48 (4.6)	13.00	.747	84
	Posttest	47.00 (5.0)	16.22			44.74 (4.9)	15.25		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

TABLE K-1. (Continued)

E. Grade 8

Site		Experimental				Control			
		Mean	SD[(b)]	r[(c)]	N[(a)]	Mean	SD[(b)]	r[(c)]	N[(a)]
Anchorage	Pretest	39.11 (5.7)	14.56	.735	37	50.55 (7.1)	15.74	.853	58
	Posttest	52.86 (7.3)	17.79			56.34 (7.6)	14.65		
Athens	Pretest	24.84 (3.9)	6.58	.375	51	28.76 (4.4)	10.62	.785	67
	Posttest	28.57 (4.4)	8.96			34.19 (5.0)	11.32		
Bronx	Pretest	23.56 (3.8)	7.45	.560	63	22.62 (3.7)	6.89	.650	16
	Posttest	31.14 (4.6)	11.84			32.25 (4.8)	9.06		
Dallas	Pretest	19.81 (3.3)	7.80	.516	47	25.30 (3.9)	7.98	.751	61
	Posttest	28.62 (4.4)	10.03			29.[illegible]0 (4.4)	9.60		
Fresno	Pretest	27.88 (4.3)	6.82	.460	83	34.15 (5.0)	14.16	.830	46
	Posttest	32.08 (4.8)	8.57			41.54 (6.2)	16.27		
Grand Rapids	Pretest	26.82 (4.2)	10.04	.726	49	35.34 (5.2)	12.27	.793	74
	Posttest	35.77 (5.3)	12.10			37.74 (5.6)	13.23		
Hammond	Pretest	34.37 (5.0)	10.30	.665	78	45.23 (6.4)	18.56	.908	97
	Posttest	42.18 (6.2)	12.68			55.18 (7.4)	17.43		
Hartford	Pretest	33.13 (4.9)	14.24	.825	82	28.66 (4.4)	13.26	.801	71
	Posttest	41.01 (6.0)	14.26			37.51 (5.6)	14.96		
Jacksonville	Pretest	28.41 (4.3)	8.56	.745	75	30.79 (4.6)	11.57	.801	75
	Posttest	34.38 (5.0)	11.90			35.01 (5.2)	14.75		
Las Vegas	Pretest	39.65 (5.8)	11.72	.745	54	37.43 (5.5)	11.89	.791	35
	Posttest	46.31 (6.6)	14.67			41.83 (6.2)	13.89		
McComb	Pretest	25.03 (3.9)	6.76	.574	70	30.60 (4.6)	11.75	.757	47
	Posttest	33.47 (4.9)	11.38			41.15 (6.0)	15.78		
Philadelphia	Pretest	21.62 (3.5)	6.75	.206	55	30.72 (4.6)	16.34	.879	47
	Posttest	20.91 (3.4)	8.04			35.87 (5.3)	17.25		
Portland	Pretest	33.93 (5.0)	9.16	.532	75	41.21 (6.0)	12.27	.793	91
	Posttest	42.55 (6.2)	13.08			51.07 (7.1)	12.88		
Rockland	Pretest	37.75 (5.6)	12.19	.786	73	58.27 (8.0)	19.84	.912	60
	Posttest	46.49 (6.6)	13.94			66.97 (9.3)	15.72		
Seattle	Pretest	38.46 (5.6)	14.00	.637	67	44.13 (6.4)	13.78	.828	77
	Posttest	41.34 (6.0)	12.80			56.13 (7.6)	15.22		
Selmer	Pretest	47.11 (6.7)	18.43	.937	80	50.82 (7.1)	15.75	.702	50
	Posttest	53.56 (7.4)	18.96			56.88 (7.8)	16.89		
Taft	Pretest	27.21 (4.2)	6.54	.564	47	32.37 (4.8)	8.63	.701	67
	Posttest	33.19 (4.9)	9.86			38.31 (5.6)	11.78		
Wichita	Pretest	35.95 (5.3)	12.86	.702	65	41.73 (6.2)	13.02	.749	66
	Posttest	42.43 (6.2)	13.98			49.05 (6.9)	15.75		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

TABLE K-1. (Continued)

F. Grade 9

Site		Experimental				Control			
		Mean	SD[b]	r[c]	N[a]	Mean	SD[b]	r[c]	N[a]
Anchorage	Pretest	50.32(7.1)	15.80	.861	40	59.11(8.0)	16.65	.848	45
	Posttest	56.27(7.6)	17.40			60.53(8.2)	20.47		
Athens	Pretest	30.30(4.5)	9.64	.831	63	37.47(5.5)	18.44	.890	66
	Posttest	37.87(5.6)	15.15			42.03(6.2)	18.99		
Bronx[d]	Pretest	--	--	--	--	--	--	--	--
	Posttest	--	--			--	--		
Dallas	Pretest	31.93(4.8)	10.37	.607	60	30.18(4.5)	14.00	.881	77
	Posttest	36.75(5.5)	10.75			35.75(5.3)	15.20		
Fresno	Pretest	28.79(4.4)	10.23	.734	63	34.20(5.0)	10.61	.537	44
	Posttest	36.95(5.5)	13.50			38.52(5.7)	13.11		
Grand Rapids	Pretest	33.97(5.0)	10.37	.714	55	39.27(5.7)	10.29	.725	63
	Posttest	40.63(6.0)	13.72			42.49(6.2)	12.19		
Hammond	Pretest	45.74(6.6)	13.96	.877	73	57.23(7.8)	16.04	.860	94
	Posttest	52.65(7.3)	14.49			63.22(8.4)	15.38		
Hartford	Pretest	30.10(4.5)	11.10	.793	70	34.50(5.2)	17.98	.857	38
	Posttest	37.77(5.6)	12.50			41.34(6.0)	17.23		
Jacksonville	Pretest	33.00(4.9)	12.61	.766	77	32.94(4.9)	13.56	.863	72
	Posttest	41.19(6.0)	14.35			39.32(5.7)	14.99		
Las Vegas	Pretest	45.60(6.6)	15.11	.893	40	52.89(7.3)	18.35	.814	44
	Posttest	54.35(7.4)	17.64			55.30(7.4)	20.28		
McComb	Pretest	30.54(4.6)	12.18	.754	71	39.98(5.8)	18.41	.792	62
	Posttest	36.83(5.5)	14.37			46.74(6.7)	18.08		
Philadelphia	Pretest	23.06(3.7)	6.34	.464	49	23.61(3.8)	10.72	.764	36
	Posttest	23.39(3.7)	7.17			27.58(4.3)	11.82		
Portland	Pretest	44.31(6.4)	14.41	.778	70	48.97(6.9)	13.88	.712	91
	Posttest	51.17(7.1)	12.94			57.26(7.8)	12.44		
Rockland	Pretest	48.75(6.9)	17.47	.890	69	69.32(9.8)	19.48	.873	69
	Posttest	57.75(8.0)	16.70			72.96(10.7)	17.02		
Seattle	Pretest	45.47(6.4)	15.45	.807	62	51.19(7.1)	13.93	.720	75
	Posttest	51.69(7.3)	16.11			60.23(8.2)	14.89		
Selmer	Pretest	44.35(6.4)	16.18	.910	86	54.62(7.4)	16.06	.837	39
	Posttest	51.51(7.3)	17.22			59.00(8.0)	16.09		
Taft	Pretest	35.57(5.3)	14.52	.830	42	37.48(5.5)	20.04	.943	25
	Posttest	44.24(6.4)	14.35			38.32(5.6)	18.79		
Wichita	Pretest	44.76(6.4)	14.27	.795	62	51.82(7.3)	18.18	.863	56
	Posttest	49.48(6.9)	15.55			56.88(7.8)	17.26		

[a] The number of students is indicated in the column labelled "N".

[b] SD = standard deviation

[c] r = pre-post linear correlation coefficient

[d] At Bronx, there were no students who had both pre- and posttest scores.

TABLE K-2. MEANS, STANDARD DEVIATIONS AND CORRELATIONS BETWEEN PRE- AND POSTTEST SCORES IN MATHEMATICS FOR FULL-YEAR STUDENTS WITH BOTH PRE- AND POSTTEST SCORES BY SITE AND GROUP

A. GRADE 1

		Experimental				Control			
Site		Mean	SD[b]	r[c]	N[a]	Mean	SD[b]	r[c]	N[a]
Anchorage	Pretest	74.46 (3)[e]	12.05	.469	46	91.04 (4)[e]	16.11	.303	53
	Posttest	36.54 (1.2)	12.82			43.40 (1.5)	11.68		
Athens	Pretest	55.63 (2)	18.81	.577	63	71.94 (3)	21.65	.724	67
	Posttest	36.95 (1.2)	12.66			41.85 (1.5)	15.10		
Bronx[d]	Pretest	--	--	--	--	--	--	--	--
	Posttest	--	--			--	--		
Dallas	Pretest	44.32 (1)	18.19	.534	41	45.40 (1)	18.99	.519	48
	Posttest	34.02 (1.0)	13.50			28.92 (.6)	11.63		
Fresno	Pretest	82.97 (4)	17.05	.363	60	83.12 (4)	15.96	.480	73
	Posttest	57.53 (2.2)	11.94			42.27 (1.5)	13.74		
Grand Rapids	Pretest	60.10 (2)	22.49	.652	61	74.76 (3)	19.73	.436	66
	Posttest	29.39 (.6)	10.64			43.11 (1.5)	14.56		
Hammond	Pretest	74.25 (3)	18.42	.657	75	81.29 (3)	17.90	.531	95
	Posttest	43.12 (1.5)	14.51			48.19 (1.7)	13.27		
Hartford	Pretest	59.90 (2)	21.74	.707	40	61.55 (2)	20.23	.626	56
	Posttest	29.38 (.6)	12.98			38.95 (1.3)	13.69		
Jacksonville	Pretest	65.08 (2)	19.21	.723	89	57.20 (2)	17.42	.614	71
	Posttest	39.97 (1.3)	12.80			22.41 (<.6)	11.47		
Las Vegas	Pretest	60.90 (2)	18.86	.646	40	76.56 (3)	17.44	.476	39
	Posttest	24.12 (<.6)	9.94			42.79 (1.5)	11.40		
McComb	Pretest	50.09 (1)	17.89	.362	35	77.11 (3)	24.81	.757	47
	Posttest	41.14 (1.4)	9.89			42.06 (1.5)	16.96		
Philadelphia	Pretest	63.49 (2)	17.81	.472	74	53.22 (2)	12.84	.494	51
	Posttest	28.49 (<.6)	7.57			34.76 (1.1)	11.73		
Portland	Pretest	79.63 (3)	20.01	.737	67	88.00 (4)	14.82	.521	64
	Posttest	38.60 (1.3)	15.81			57.22 (2.1)	11.70		
Rockland	Pretest	85.26 (4)	14.45	.504	83	93.71 (5)	15.48	.616	73
	Posttest	49.58 (1.8)	13.20			56.00 (2.1)	14.21		
Seattle	Pretest	68.46 (2)	10.36	.646	13	88.36 (4)	16.94	.730	69
	Posttest	28.54 (.6)	8.99			50.07 (1.8)	19.24		
Selmer	Pretest	70.29 (2)	17.34	.683	90	73.42 (3)	21.97	.626	33
	Posttest	45.42 (1.6)	12.58			43.12 (1.5)	14.78		
Taft	Pretest	54.33 (2)	21.99	.438	51	61.53 (2)	16.61	.635	68
	Posttest	39.35 (1.3)	12.39			45.32 (1.6)	12.24		
Wichita	Pretest	73.24 (3)	12.29	.331	71	73.31 (3)	18.26	.656	55
	Posttest	36.39 (1.1)	10.49			39.31 (1.3)	11.58		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

(d) At Bronx the E group did not take the evaluation pretest

(e) Stanine values are given for Grade 1

TABLE K-2. (Continued)

B. Grade 2

Site		Experimental				Control			
		Mean	SD[b]	r[c]	N[a]	Mean	SD[b]	r[c]	N[a]
Anchorage	Pretest	27.87(1.4)	7.95	.541	46	35.32(1.5)	9.48	.755	41
	Postest	46.93(2.2)	9.18			50.15(2.4)	7.22		
Athens	Pretest	29.88(1.4)	10.20	.727	73	30.43(1.4)	10.75	.770	79
	Posttest	43.88(2.0)	10.59			44.45(2.0)	11.63		
Bronx	Pretest	21.22(1.2)	10.07	.780	27	22.48(1.2)	9.04	.447	50
	Posttest	38.81(1.7)	13.60			42.72(2.0)	7.85		
Dallas	Pretest	22.73(1.2)	11.72	.655	48	18.86(1.2)	6.60	.453	44
	Posttest	35.23(1.5)	12.99			23.98(1.3)	8.39		
Fresno	Pretest	30.61(1.4)	9.45	.653	83	32.44(1.5)	9.59	.658	71
	Posttest	42.88(2.0)	10.21			47.69(2.3)	8.16		
Grand Rapids	Pretest	33.45(1.5)	9.99	.731	53	31.70(1.5)	9.81	.640	71
	Posttest	44.57(2.1)	9.72			44.70(2.1)	10.69		
Hammond	Pretest	29.73(1.4)	10.93	.694	73	31.67(1.5)	12.34	.710	94
	Posttest	43.45(2.0)	10.21			46.03(2.2)	11.70		
Hartford	Pretest	18.96(1.2)	7.74	.569	49	26.74(1.3)	10.16	.537	57
	Posttest	34.39(1.5)	9.12			42.68(2.0)	10.51		
Jacksonville	Pretest	22.60(1.2)	8.33	.731	88	21.20(1.2)	7.48	.693	61
	Posttest	43.25(2.0)	10.73			36.07(1.6)	10.42		
Las Vegas	Pretest	18.46(1.1)	9.04	.526	41	22.59(1.2)	7.68	.595	39
	Posttest	27.46(1.3)	10.86			37.62(1.7)	11.05		
McComb	Pretest	19.12(1.2)	12.87	.535	49	26.90(1.3)	8.73	.762	40
	Posttest	37.02(1.6)	9.97			43.72(2.0)	9.64		
Philadelphia	Pretest	26.22(1.3)	10.56	.787	79	20.02(1.2)	8.33	.533	58
	Posttest	37.84(1.7)	10.73			37.36(1.6)	11.63		
Portland	Pretest	38.20(1.7)	10.60	.769	85	39.31(1.7)	9.61	.762	81
	Posttest	48.69(2.3)	7.77			52.81(2.6)	7.72		
Rockland	Pretest	33.14(1.5)	10.82	.702	9	44.17(2.0)	9.39	.599	65
	Posttest	49.63(2.4)	9.06			57.52(3.5)	4.32		
Seattle	Pretest	28.81(1.4)	10.37	.647	63	35.26(1.5)	10.09	.713	61
	Posttest	42.76(2.0)	12.87			52.54(2.6)	7.56		
Selmer	Pretest	34.77(1.5)	11.51	.787	91	36.52(1.6)	10.59	.729	42
	Posttest	48.43(2.3)	10.43			47.76(2.3)	9.95		
Taft	Pretest	25.78(1.3)	7.55	.672	60	30.45(1.4)	10.23	.698	84
	Posttest	41.17(1.9)	7.05			45.13(2.1)	9.84		
Wichita	Pretest	27.48(1.3)	11.56	.624	65	32.21(1.5)	13.29	.650	68
	Posttest	39.82(1.8)	12.41			43.81(2.0)	11.74		

[a] The number of students is indicated in the column labelled "N".

[b] SD = standard deviation

[c] r = pre-post linear correlation coefficient

TABLE K-2. (Continued)

C. Grade 3

		Experimental				Control			
Site		Mean	SD[b]	r[c]	N[a]	Mean	SD[b]	r[c]	N[a]
Anchorage	Pretest	46.33(2.2)	9.46	.574	52	57.90(2.4)	15.96	.824	42
	Posttest	67.96(2.9)	13.94			78.81(3.3)	15.81		
Athens	Pretest	45.80(2.2)	14.78	.702	83	54.81(2.4)	17.34	.848	69
	Posttest	64.61(2.7)	17.08			68.26(2.9)	20.34		
Bronx	Pretest	45.47(2.2)	11.78	.692	40	39.51(1.9)	10.89	.666	57
	Posttest	68.65(3.0)	18.86			67.19(2.8)	13.19		
Dallas	Pretest	35.09(1.6)	12.14	.397	47	35.33(1.6)	10.71	.544	61
	Posttest	58.83(2.5)	18.50			39.36(1.9)	12.77		
Fresno	Pretest	47.44(2.2)	11.50	.752	73	53.59(2.4)	15.16	.793	80
	Posttest	59.03(2.5)	15.86			70.47(3.0)	16.69		
Grand Rapids	Pretest	40.64(2.0)	13.45	.687	78	44.50(2.2)	18.89	.7[illegible]4	62
	Posttest	57.62(2.4)	15.97			62.90(2.6)	18.76		
Hammond	Pretest	46.68(2.2)	15.02	.794	73	56.63(2.4)	18.24	.853	107
	Posttest	60.88(2.6)	19.02			73.74(3.1)	17.86		
Hartford	Pretest	41.02(2.0)	14.51	.645	48	47.60(2.3)	16.52	.726	48
	Posttest	48.92(2.3)	13.64			66.19(2.8)	17.72		
Jacksonville	Pretest	38.04(1.8)	12.49	.799	92	36.91(1.7)	11.29	.673	75
	Posttest	58.84(2.5)	19.10			50.51(2.3)	13.78		
Las Vegas	Pretest	27.74(1.3)	8.83	.486	64	30.03(1.4)	8.03	.447	25
	Posttest	38.41(1.8)	11.73			47.94(2.3)	15.82		
McComb	Pretest	35.46(1.6)	10.00	.675	69	56.62(2.4)	13.24	.804	50
	Posttest	53.35(2.3)	14.36			82.60(3.5)	17.10		
Philadelphia	Pretest	35.57(1.7)	12.12	.728	65	35.96(1.7)	10.38	.661	53
	Posttest	52.05(2.3)	16.92			53.45(2.3)	13.[illegible]3		
Portland	Pretest	53.94(2.4)	12.23	.677	81	62.57(2.6)	15.56	.732	74
	Posttest	69.59(3.0)	16.59			82.84(3.5)	15.99		
Rockland	Pretest	50.13(2.3)	15.85	.824	89	65.57(2.7)	20.06	.805	70
	Posttest	73.27(3.1)	17.63			85.64(3.5)	14.65		
Seattle	Pretest	39.92(1.9)	12.00	.735	61	64.09(2.6)	17.34	.669	75
	Posttest	57.75(2.4)	15.87			85.23(3.5)	17.47		
Selmer	Pretest	55.95(2.4)	16.55	.729	88	64.29(2.6)	14.74	.715	45
	Posttest	86.27(3.5)	14.00			79.87(3.3)	18.16		
Taft	Pretest	50.55(2.3)	11.86	.836	73	53.18(2.3)	17.17	.826	85
	Posttest	70.04(3.0)	13.37			65.85(2.8)	18.07		
Wichita	Pretest	41.38(2.0)	14.95	.759	66	40.32(1.9)	17.01	.616	37
	Posttest	54.64(2.4)	15.81			59.46(2.5)	15.33		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

TABLE K-2. (Continued)

D. Grade 7

Site		Experimental				Control			
		Mean	SD[b]	r[c]	N[a]	Mean	SD[b]	r[c]	N[a]
Anchorage	Pretest	47.25(4.9)	14.02	.653	44	57.50(5.6)	17.91	.805	62
	Posttest	58.36(5.6)	18.07			63.97(5.9)	20.57		
Athens	Pretest	33.95(4.2)	10.52	.554	61	45.50(4.8)	16.25	.752	72
	Posttest	49.31(5.1)	18.22			55.37(5.4)	20.02		
Bronx	Pretest	39.07(4.5)	13.48	.817	58	38.14(4.4)	10.71	.675	7
	Posttest	45.97(4.8)	15.65			39.71(4.5)	10.66		
Dallas	Pretest	32.53(4.1)	13.36	.808	51	27.26(3.7)	6.40	.374	42
	Posttest	36.22(4.3)	13.91			32.45(4.0)	14.08		
Fresno	Pretest	33.76(4.2)	8.22	.574	80	39.07(4.5)	10.71	.456	60
	Posttest	40.34(4.5)	10.89			54.55(5.4)	17.46		
Grand Rapids	Pretest	36.46(4.3)	13.50	.682	60	40.74(4.6)	15.61	.854	74
	Posttest	43.77(4.7)	14.32			52.66(5.3)	18.54		
Hammond	Pretest	49.00(5.1)	14.54	.774	88	59.10(5.6)	20.58	.887	92
	Posttest	60.73(5.7)	18.68			69.18(6.2)	21.79		
Hartford	Pretest	40.81(4.6)	14.73	.835	62	44.13(4.7)	17.33	.843	52
	Posttest	52.39(5.2)	16.45			55.75(5.4)	19.91		
Jacksonville	Pretest	35.38(4.2)	13.60	.760	74	40.10(4.5)	14.23	.820	71
	Posttest	43.91(4.7)	15.41			49.42(5.1)	18.49		
Las Vegas	Pretest	48.19(4.9)	15.00	.749	63	43.45(4.7)	12.51	.759	42
	Posttest	58.29(5.6)	16.33			57.52(5.6)	18.62		
McComb	Pretest	40.10(4.5)	14.82	.876	69	37.96(4.4)	15.48	.864	50
	Posttest	52.50(5.3)	17.93			50.18(5.1)	19.34		
Philadelphia	Pretest	29.93(3.9)	8.25	.451	45	25.00(3.5)	8.40	.346	32
	Posttest	39.71(4.5)	10.04			34.22(4.2)	9.48		
Portland	Pretest	46.68(4.9)	10.39	.532	81	57.89(5.6)	17.54	.846	95
	Posttest	60.84(5.7)	13.43			69.64(6.2)	18.78		
Rockland	Pretest	54.07(5.3)	17.49	.846	82	72.37(6.4)	19.25	.821	62
	Posttest	65.61(6.0)	18.06			86.66(7.6)	17.61		
Seattle	Pretest	47.05(4.9)	16.44	.700	64	65.27(5.9)	19.86	.732	79
	Posttest	50.47(5.1)	20.12			80.99(7.1)	20.46		
Selmer	Pretest	58.14(5.6)	23.15	.864	81	47.78(4.9)	20.05	.793	50
	Posttest	64.69(5.9)	16.30			53.27(5.3)	18.79		
Taft	Pretest	39.28(4.5)	13.21	.782	60	52.71(5.3)	18.78	.894	65
	Posttest	46.02(4.8)	15.54			59.75(5.7)	20.45		
Wichita	Pretest	41.07(4.6)	14.90	.803	74	38.37(4.4)	13.08	.715	82
	Posttest	48.55(5.1)	16.62			46.22(4.8)	15.48		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

TABLE K-2. (Continued)

E. Grade 8

Site		Experimental				Control			
		Mean	SD(b)	r(c)	N(a)	Mean	SD(b)	r(c)	N(a)
Anchorage	Pretest	41.32(5.6)	16.76	.760	38	48.14(6.2)	12.53	.768	58
	Posttest	47.97(6.4)	17.34			57.76(7.1)	17.84		
Athens	Pretest	31.42(4.5)	8.99	.422	45	31.67(4.6)	10.19	.643	63
	Posttest	34.53(5.2)	11.12			36.98(5.4)	12.11		
Bronx	Pretest	30.30(4.4)	9.99	.665	23	30.17(4.4)	7.79	.385	12
	Posttest	35.22(5.2)	9.43			37.83(5.6)	6.89		
Dallas	Pretest	21.41(3.7)	7.15	.287	28	28.20(4.3)	8.63	.693	61
	Posttest	35.19(5.2)	8.42			35.25(5.2)	10.70		
Fresno	Pretest	30.79(4.5)	8.84	.442	80	39.83(5.4)	19.53	.845	42
	Posttest	35.41(5.2)	9.19			47.62(6.4)	20.63		
Grand Rapids	Pretest	34.68(4.9)	11.45	.710	50	38.63(5.4)	12.75	.729	79
	Posttest	37.31(5.4)	12.16			47.87(6.4)	15.38		
Hammond	Pretest	39.83(5.4)	11.94	.679	77	50.92(6.4)	17.37	.893	93
	Posttest	46.85(6.4)	12.74			60.74(7.5)	20.07		
Hartford	Pretest	41.24(5.6)	13.83	.854	70	36.37(5.1)	12.49	.879	73
	Posttest	45.07(6.2)	14.40			42.30(5.9)	15.42		
Jacksonville	Pretest	35.85(5.1)	10.85	.724	75	39.80(5.4)	11.45	.788	74
	Posttest	41.72(5.9)	12.67			45.74(6.2)	15.32		
Las Vegas	Pretest	40.06(5.4)	12.30	.772	51	42.09(5.7)	11.65	.536	33
	Posttest	44.82(6.2)	17.03			48.42(6.4)	17.01		
McComb	Pretest	36.03(5.1)	12.02	.846	69	44.23(5.9)	14.02	.867	47
	Posttest	43.72(6.0)	15.17			55.02(6.9)	18.14		
Philadelphia	Pretest	25.89(4.1)	7.58	.556	53	39.56(5.4)	18.38	.900	43
	Posttest	31.89(4.8)	7.50			44.12(6.0)	19.47		
Portland	Pretest	43.71(5.9)	12.82	.729	72	49.85(6.4)	15.87	.733	89
	Posttest	49.46(6.6)	14.64			61.93(7.5)	18.22		
Rockland	Pretest	46.61(6.0)	14.94	.860	71	68.12(7.8)	20.47	.896	60
	Posttest	56.25(6.9)	17.84			79.72(9.2)	18.58		
Seattle	Pretest	43.94(5.9)	12.52	.626	68	50.24(6.4)	16.48	.813	75
	Posttest	42.09(5.9)	14.23			67.35(8.0)	19.94		
Selmer	Pretest	57.37(6.8)	20.35	.952	78	63.61(7.5)	16.34	.721	51
	Posttest	66.92(8.0)	22.26			71.43(8.2)	19.07		
Taft	Pretest	34.35(4.8)	9.	.678	48	43.13(5.7)	14.49	.696	68
	Posttest	40.52(5.9)	11.64			48.25(6.4)	16.48		
Wichita	Pretest	35.39(4.9)	10.00	.455	69	43.37(5.7)	14.13	.575	67
	Posttest	44.30(6.0)	11.75			50.91(6.7)	15.47		

(a) The number of students is indicated in the column labelled "N".

(b) SD = standard deviation

(c) r = pre-post linear correlation coefficient

TABLE K-2. (Continued)

F. Grade 9

Site		Experimental				Control			
		Mean	SD(b)	r(c)	N(a)	Mean	SD(b)	r(c)	N(a)
Anchorage	Pretest	52.18(6.6)	15.41	.786	45	53.50(6.7)	20.28	.882	40
	Posttest	64.91(7.8)	18.30			55.67(6.9)	20.38		
Athens	Pretest	37.41(5.2)	10.91	.664	64	43.86(5.9)	18.37	.914	64
	Posttest	49.30(6.6)	18.68			50.91(6.7)	20.60		
Bronx(d)	Pretest	--	--	--	--	--	--	--	--
	Posttest	--	--			--	--		
Dallas	Pretest	36.94(5.2)	9.93	.601	55	39.50(5.4)	12.54	.813	76
	Posttest	43.58(6.0)	13.07			45.28(6.2)	13.87		
Fresno	Pretest	33.93(4.8)	12.13	.620	60	38.37(5.3)	13.66	.759	48
	Posttest	40.29(5.7)	14.38			42.87(6.0)	16.41		
Grand Rapids	Pretest	41.31(5.6)	12.89	.741	58	44.33(5.9)	12.82	.744	58
	Posttest	45.73(6.2)	15.22			49.29(6.6)	15.83		
Hammond	Pretest	51.93(6.6)	16.55	.864	73	63.95(7.5)	17.61	.860	93
	Posttest	63.43(7.6)	18,86			75.74(8.8)	15.88		
Hartford	Pretest	37.82(5.3)	13.16	.794	68	46.53(6.0)	19.75	.892	34
	Posttest	44.78(6.2)	15.53			49.18(6.6)	19.26		
Jacksonville	Pretest	41.52(5.7)	13.61	.740	73	42.95(5.7)	13.57	.810	73
	Posttest	48.32(6.4)	15.20			48.66(6.6)	16.65		
Las Vegas	Pretest	51.31(6.4)	16.28	.483	35	55.07(6.7)	16.89	.871	46
	Posttest	60.97(7.5)	17.39			63.22(7.6)	18.35		
McComb	Pretest	41.72(5.7)	12.58	.792	71	55.88(6.8)	19.80	.910	60
	Posttest	48.20(6.4)	15.55			63.37(7.6)	21.38		
Philadelphia	Pretest	29.85(4.4)	8.90	.633	53	32.24(4.6)	10.11	.630	33
	Posttest	32.68(4.9)	8.40			37.82(5.6)	10.59		
Portland	Pretest	52.84(6.6)	13.96	.772	70	55.55(6.8)	15.44	.594	89
	Posttest	60.19(7.3)	14.49			67.62(8.0)	15.11		
Rockland	Pretest	55.48(6.7)	19.90	.846	73	85.30(9.2)	19.45	.910	67
	Posttest	65.26(7.8)	20.49			86.51(10.4)	23.28		
Seattle	Pretest	54.92(6.7)	17.14	.795	61	53.27(6.6)	16.61	.719	75
	Posttest	60.66(7.5)	19.05			70.76(8.2)	20.48		
Selmer	Pretest	55.89(6.8)	17.13	.880	87	66.92(7.6)	20.85	.879	40
	Posttest	65.40(7.8)	17.77			79.40(9.0)	19.56		
Taft	Pretest	51.17(6.4)	13.91	.908	42	42.17(5.7)	18.07	.865	23
	Posttest	58.43(7.1)	15.82			45.48(6.2)	21.01		
Wichita	Pretest	43.53(5.9)	13.84	.613	70	52.73(6.6)	20.09	.752	56
	Posttest	51.80(6.7)	17.18			62.07(7.5)	19.94		

(a) The number of students is indicated in the column labelled "N".

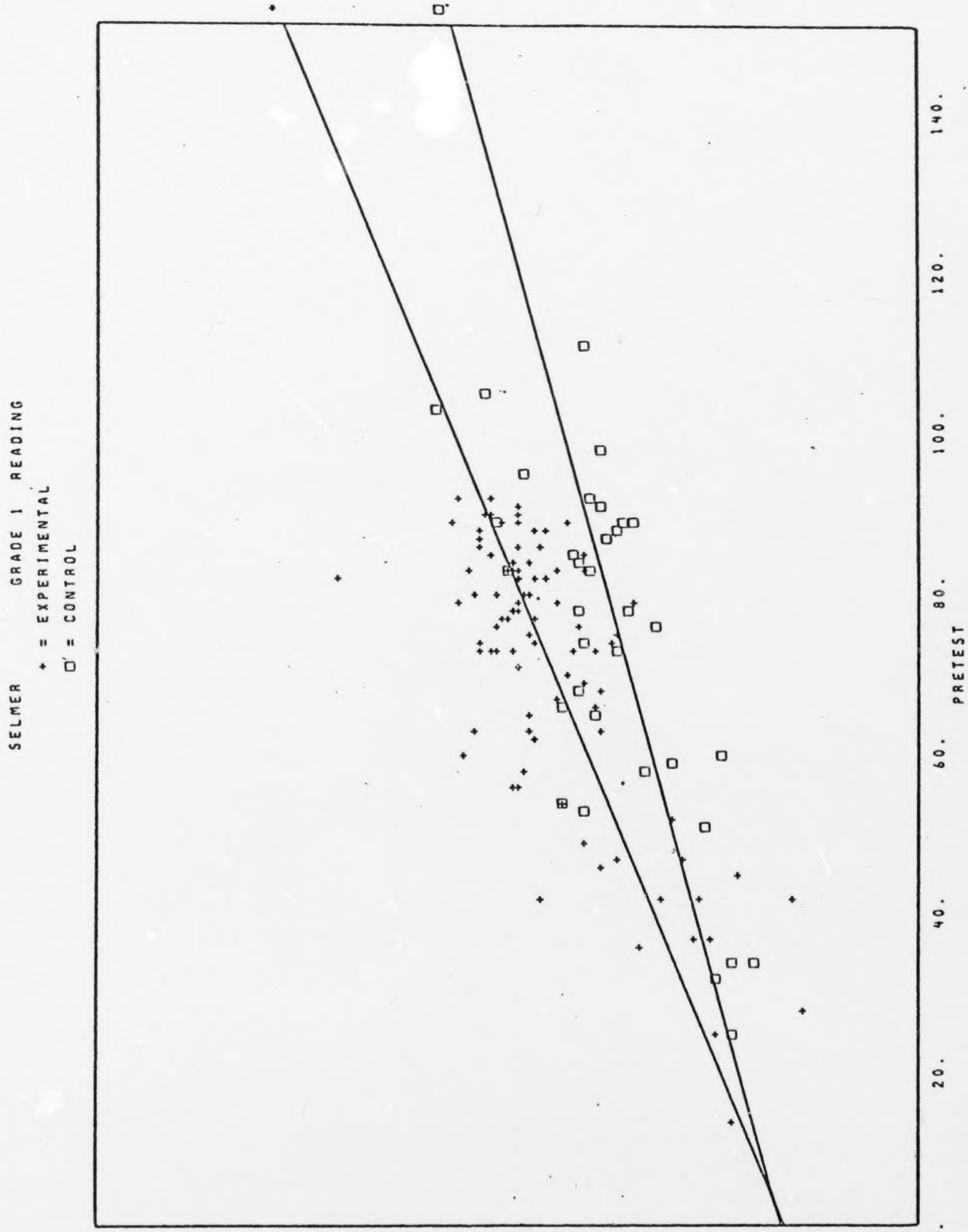
(b) SD = standard deviation

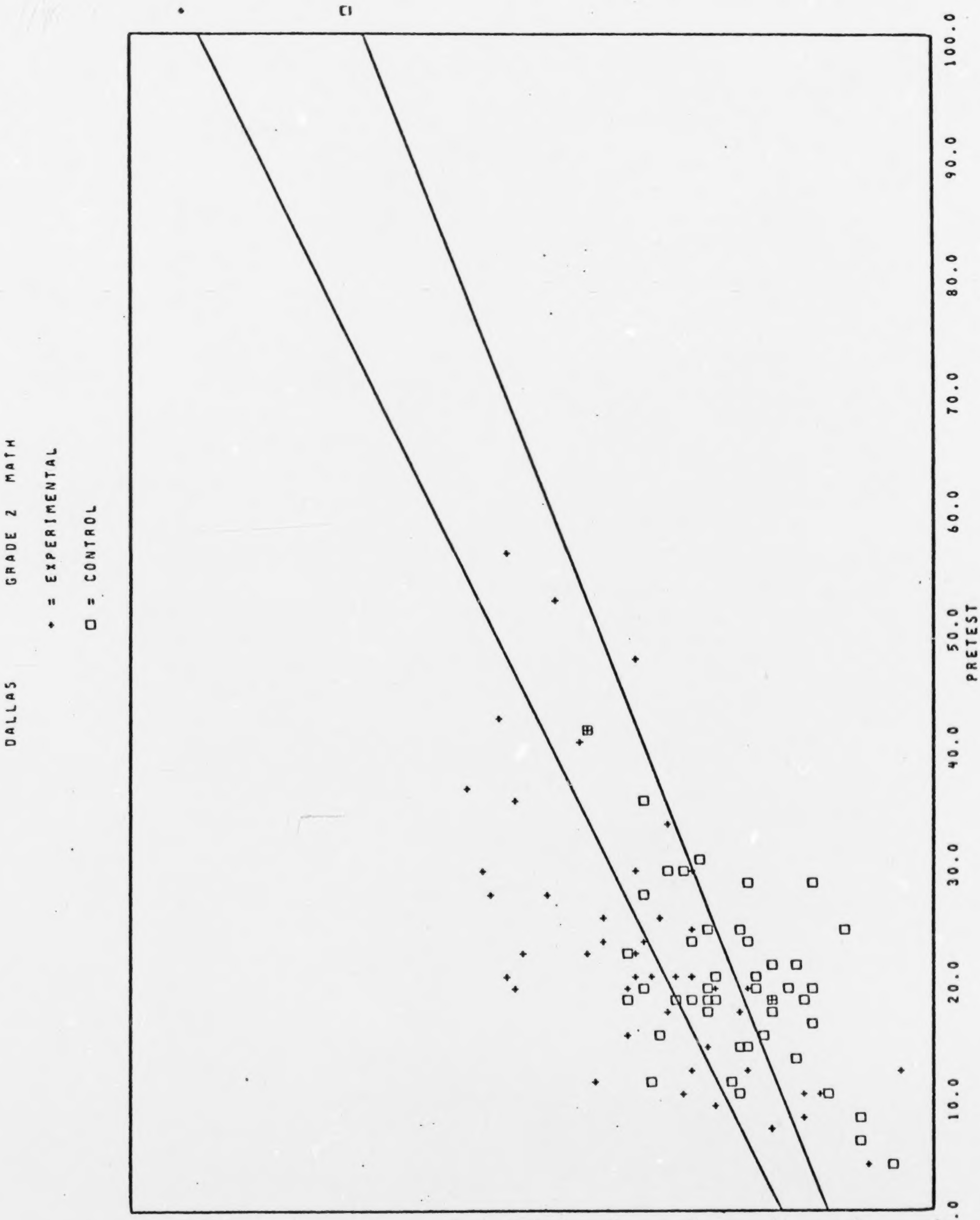
(c) r = pre-post linear correlation coefficient

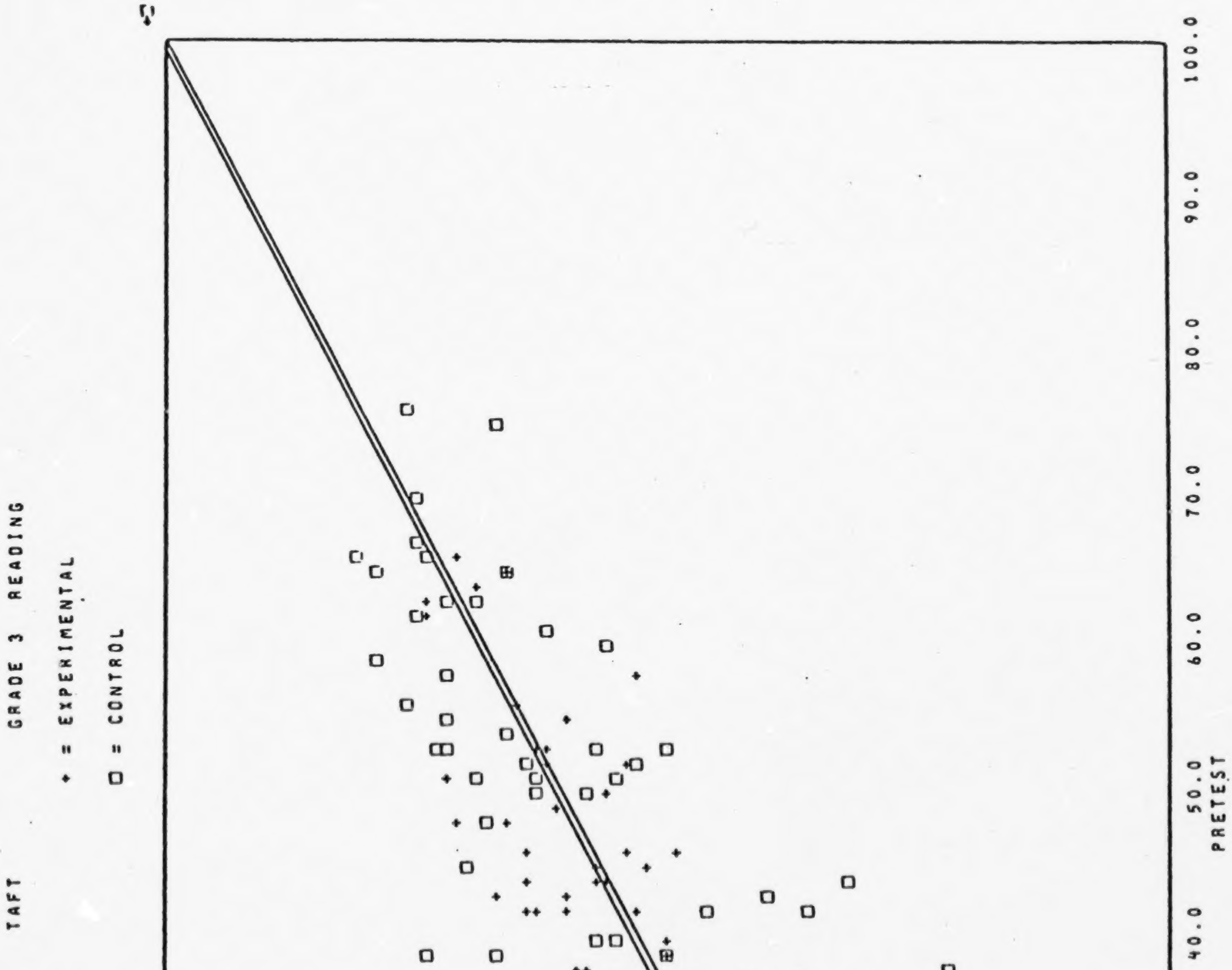
(d) At Bronx, there were no students who had both pre- and posttest scores.

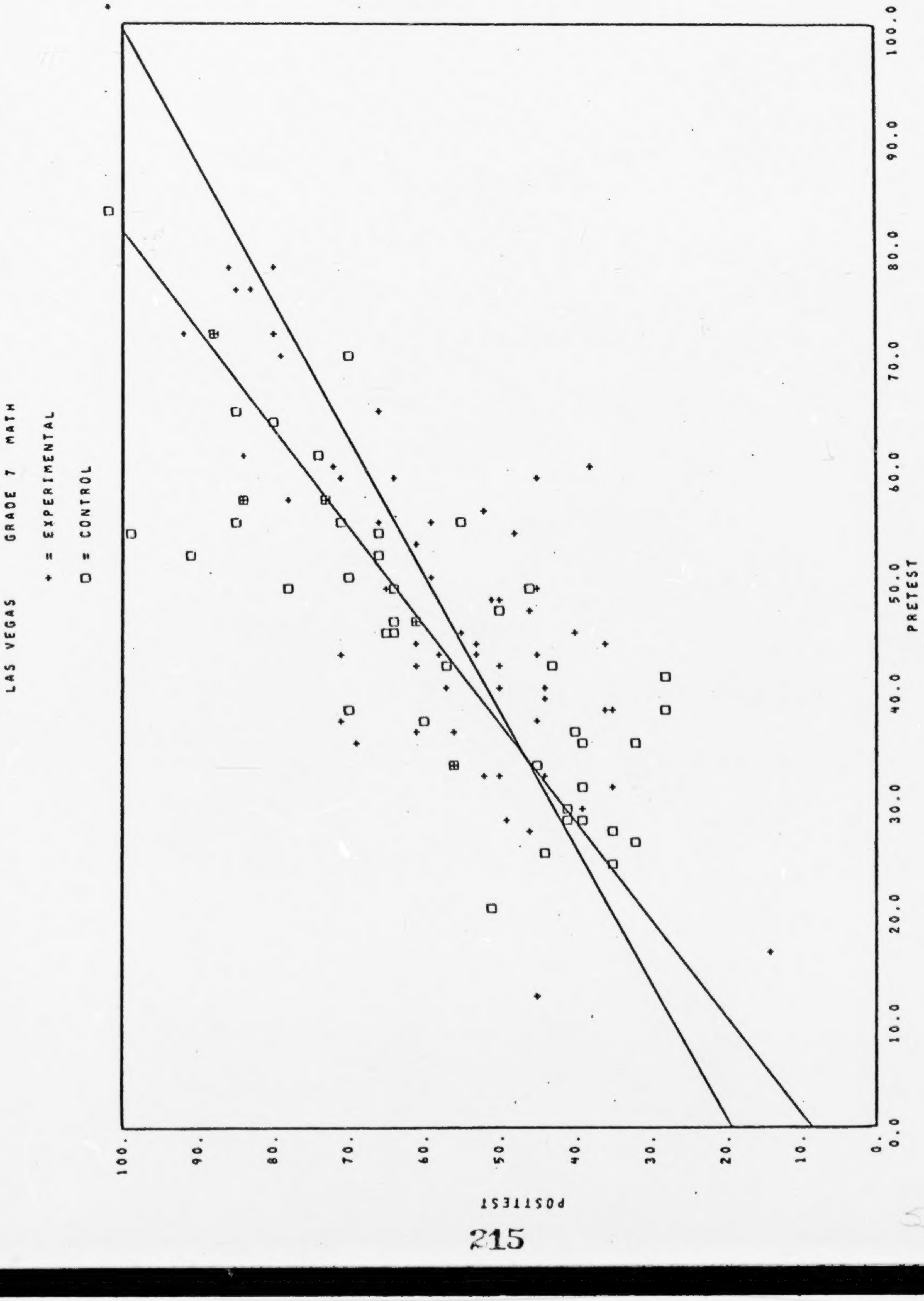
APPENDIX L

ILLUSTRATIVE SCATTERPLOTS
(IN RAW SCORE UNITS)

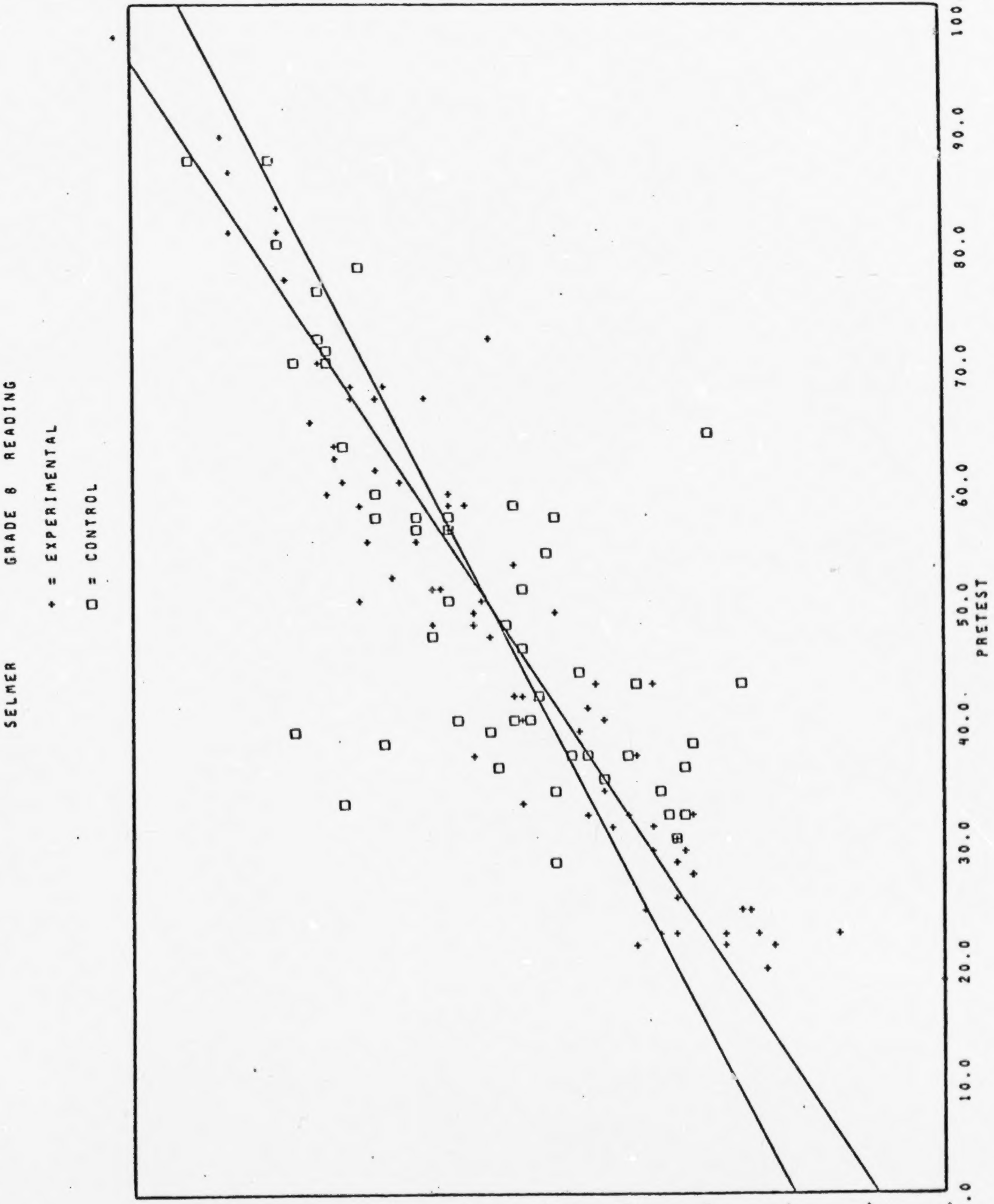

SELMER GRADE 1 READING
+ = EXPERIMENTAL
□ = CONTROL
150.
125.
100.
75.
50.
25.
0.
POSTTEST
0.
20.
40.
60.
80.
100.
120.
140.
PRETEST

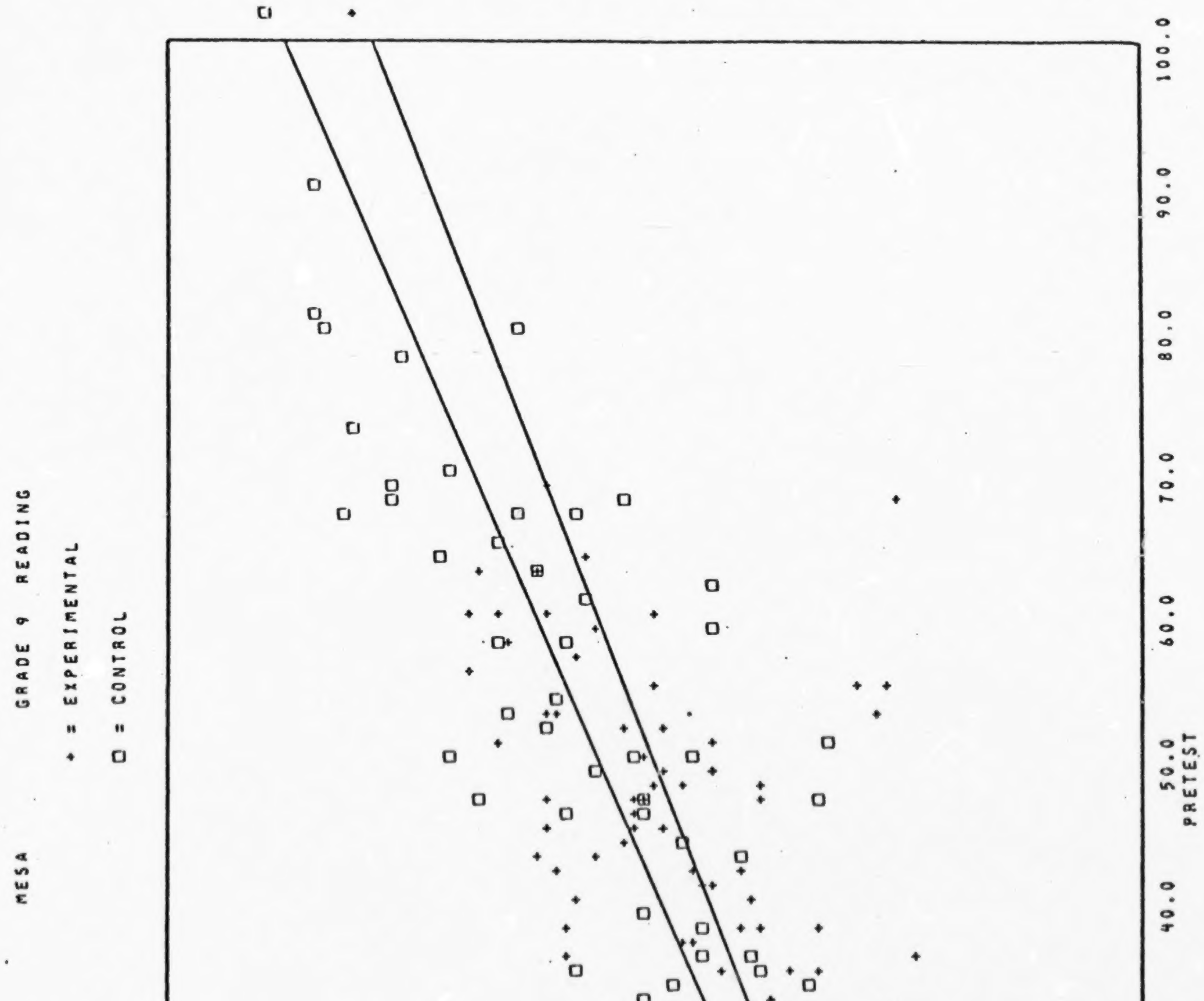

DALLAS
GRADE 2 MATH
+ = EXPERIMENTAL
□ = CONTROL
PRETEST
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0
POSTTEST
0.0
10.
20.
30.
40.
50.
60.
70.
80.
90.
100.


TAFT
GRADE 3 READING
+ = EXPERIMENTAL
□ = CONTROL
PRETEST
POSTTEST
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0
0.0
10.
20.
30.
40.
50.
60.
70.
80.
90.
100.



LAS VEGAS GRADE 7 MATH


SELMER
GRADE 8 READING
+ = EXPERIMENTAL
□ = CONTROL
PRETEST
POSTTEST
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0
0.
10.
20.
30.
40.
50.
60.
70.
80.
90.
100.


MESA
GRADE 9 READING
+ = EXPERIMENTAL
□ = CONTROL
PRETEST
POSTTEST
0.0
10.0
20.0
30.0
40.0
50.0
60.0
70.0
80.0
90.0
100.0
0.
10.
20.
30.
40.
50.
60.
70.
80.
90.
100.

55

APPENDIX M

PRE-POST REGRESSIONS HAVING SIGNIFICANT GROUP-BY-PRETEST INTERACTIONS

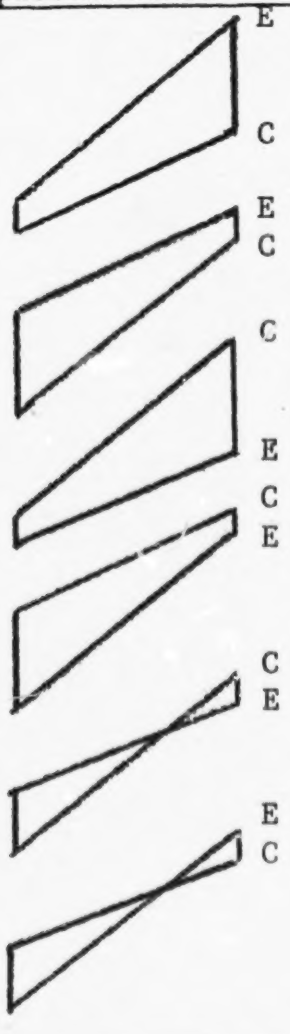
The pre-post regression model fitted to the data at each grade/site/subject combination includes a group-by-pretest interaction term. The inclusion of this term permits the model to detect differences in the slopes of the E and C posttest regression lines as functions of pretest scores. If such a difference in slopes exists, then the inclusion of this term will result in an improved estimate of the residual error. This improved estimate, in turn, will yield a better statistical assessment of the difference between the E and C regression lines evaluated at the mean of the pretest scores for the combined E and C groups. This latter assessment is taken to be the primary measure of the impact of the experiment at each grade/site/subject combination.

The inclusion of the interaction term in the model also yields secondary benefits. Whenever the interaction is significant, the regression lines will have markedly different slopes, and may even intersect in the range of the pretest scores. This has the effect of showing reversed differences between the fitted E and C posttest means corresponding to low and high pretest scores. Because of the importance of this kind of information, a summary is given of the results obtained at all those grade/site/subject combinations which yielded a group-by-pretest interaction coefficient having a t-value greater than 2.0.

Table M-1 shows a convenient classification scheme for the six types of E and C regression plots that can result. These types are numbered 1 through 6. The table shows, for example, that the type 1 plot consists of having the posttest regression line for the E group above that of the C group for the entire range of pretest scores common to the two groups. In addition, the type 1 plot yields regression lines that diverge for the larger pretest scores. Types 5 and 6 illustrate those interactions in which a reversal occurs at the two ends of the range of pretest scores. The results showing significant interactions are classified according to the types shown in this table.

56/57

TABLE M-1. DEFINITIONS OF TYPES OF REGRESSION PLOTS FOR POSTTEST VERSUS PRETEST FOR E AND C GROUPS



Type Number	Graphical Representation(1)	Description
1		Posttest regression line for E group is above that for C group and the lines diverge for higher pretest scores.
2		Same as Type 1 except that lines converge for higher pretest scores.
3		Same as Type 1 except that line for C group is above that for E group.
4		Same as Type 2 except that line for C group is above that for E group.
5		Posttest regression lines intersect within range of pretest scores with line for E group above that for C group at low pretest scores.
6		Same as Type 5 except that line for C group is above that for E group at low pretest scores.

(1) These representations are intended to depict the general form of the posttest regression lines for the experimental (E) and control (C) groups over the range of pretest scores common to the two groups.

Table M-2 shows the results of these classifications. The symbols R and M denote reading and mathematics, respectively. The plot types are shown in column 4. For types 5 and 6, which have intersecting regression lines, the last column shows the approximate percentage of the pretest range for which the regression line for the E group is above that of the C group. In grade 1 at McComb, for example, the regression lines for reading intersects to give a type 5 plot. This plot shows that the lower 45 percent of the pretest range has an associated E group posttest regression line that is higher than that of the C group. In grade 3 at Selmer, the plot is of type 2 for reading. This means that regression lines have different slopes and intersect beyond the upper limit of the pretest range. In this case, the regression line for the E group is higher than that of the C group for 100 percent of the pretest range.

An examination of this table shows that 40 interactions occur. The frequencies of occurrence for the six plot types are seen to be given by, 3, 3, 3, 0, 17, 14, respectively, for types 1 through 6. The two types of intersecting plots, 5 and 6, occur with approximately the same frequency.

TABLE M-2. CLASSIFICATION OF PRE-POST REGRESSION PLOTS HAVING SIGNIFICANT GROUP-BY-PRETEST INTERACTIONS

Grade	Site	Subject	Plot Type(1)	Pretest Range for Which E>C, Percent
1	Athens	R	5	Lower 58
	Fresno	M	3	0
	Philadelphia	M	3	0
	Taft	M	5	Lower 38
	McComb	R	5	Lower 45
		M	5	Lower 80
	Portland	R	6	Upper 12
	Jacksonville	R	1	100
2	Selmer	R	5	Lower 72
	Rockland	R	6	Upper 27
		M	6	Upper 21
	McComb	M	5	Lower 16
	Seattle	R	6	Upper 20
	Portland	R	6	Upper 29
	Jacksonville	M	6	Upper 80
	Bronx	M	6	Upper 45
3	Selmer	R	2	100
		M	2	100
	Rockland	R	1	100
		M	6	Upper 39
	Philadelphia	R	6	Upper 39
	McComb	R	6	Upper 12
	Jacksnoville	M	1	100
7	Grand Rapids	M	5	Lower 9
8	Selmer	R	6	Upper 63
	Dallas	M	5	Lower 52
	Rockland	M	6	Upper 23
	Fresno	R	5	Lower 21
		M	5	Lower 19
	Philadelphia	R	5	Lower 10
		M	5	Lower 44
	Hartford	M	5	Lower 20
	Seattle	R	3	0
	Hammond	M	5	Lower 19
9	Athens	R	6	Upper 76
	Dallas	R	5	Lower 37
	Rockland	M	2	100
	Las Vegas	M	5	Lower 49
	Grand Rapids	M	5	Lower 35
	Hammond	M	6	Upper 36

(1) Plot types correspond to those given in Table M-1.

APPENDIX N

SAMPLE SIZE FOR PRE-POST MODEL REGRESSION ANALYSES

61

TABLE N-1. NUMBER OF EXPERIMENTAL (E), CONTROL (C), AND COMPARISON (R) FULL-YEAR STUDENTS HAVING BOTH A PRETEST AND POSTTEST SCORE, BY SITE, GRADE, AND SUBJECT AREA

Site	Reading			Mathematics		
	E	C	R	E	C	R
A. Grade 1						
Selmer	91	33	-	90	33	-
Athens	64	67	-	63	67	-
Wichita	71	56	34	71	55	34
Dallas	40	46	-	41	48	-
Anchorage	46	52	-	46	53	-
Rockland	82	75	76	83	73	76
Las Vegas	44	41	-	40	39	-
Fresno	60	76	-	60	73	-
Philadelphia	76	51	-	74	51	-
Taft	51	72	-	51	68	-
Grand Rapids	60	68	-	61	66	-
Hartford	40	57	-	40	56	-
McComb	35	47	31	35	47	31
Seattle	55	65	-	13	69	-
Portland	68	64	-	67	64	-
Jacksonville	90	72	45	89	71	45
Hammond	74	97	34	75	95	34
Bronx	-	-	-	-	-	-
B. Grade 2						
Selmer	92	41	-	91	42	-
Athens	75	85	-	73	79	-
Wichita	65	72	44	65	68	38
Dallas	73	63	-	48	44	-
Anchorage	48	44	-	46	41	-
Rockland	80	67	63	79	65	62
Las Vegas	49	38	-	41	39	-
Fresno	81	82	-	83	71	-
Philadelphia	82	65	-	79	58	-
Taft	61	89	-	60	84	-
Grand Rapids	71	72	-	53	71	-
Hartford	54	67	-	49	57	-
McComb	48	31	27	49	40	26
Seattle	69	71	-	63	61	-
Portland	84	80	-	85	81	-
Jacksonville	93	77	48	88	61	44
Hammond	75	96	59	73	94	58
Bronx	42	52	-	27	50	-

TABLE N-1. (Continued)

	Reading			Mathematics		
Site	E	C	R	E	C	R
C. Grade 3						
Selmer	88	45	-	88	45	-
Athens	85	71	-	83	69	-
Wichita	63	42	38	66	37	35
Dallas	71	76	-	47	61	-
Anchorage	50	45	-	52	42	-
Rockland	92	75	108	89	70	110
Las Vegas	68	31	-	64	25	-
Fresno	78	85	-	73	80	-
Philadelphia	66	50	-	65	53	-
Taft	73	83	-	73	85	-
Grand Rapids	81	67	-	78	62	-
Hartford	53	46	-	48	48	-
McComb	69	51	37	69	50	37
Seattle	72	79	-	61	75	-
Portland	79	77	-	81	74	-
Jacksonville	91	76	47	92	75	45
Hammond	74	111	51	73	107	49
Bronx	64	55	-	40	57	-
D. Grade 7						
Selmer	84	49	-	81	50	-
Athens	64	70	38	61	72	35
Wichita	72	84	27	74	82	29
Dallas	57	43	-	51	42	-
Anchorage	46	57	-	44	62	-
Rockland	80	59	101	82	62	99
Las Vegas	64	46	-	63	42	-
Fresno	83	64	-	80	60	-
Philadelphia	45	39	-	45	32	-
Taft	59	61	-	60	65	-
Grand Rapids	59	71	-	60	74	-
Hartford	71	58	45	62	52	32
McComb	69	49	-	69	50	-
Seattle	70	82	-	64	79	-
Portland	81	97	-	81	95	-
Jacksonville	75	71	15	74	71	17
Hammond	88	96	-	88	92	-
Bronx	44	12	-	58	7	-

TABLE N-1. (Continued)

Site	Reading			Mathematics		
	E	C	R	E	C	R
			E. Grade 8			
Selmer	80	50	-	78	51	-
Athens	51	67	26	45	63	27
Wichita	65	66	23	69	67	18
Dallas	47	61	-	28	61	-
Anchorage	37	58	-	38	58	-
Rockland	73	60	103	71	60	104
Las Vegas	54	35	-	51	33	-
Fresno	83	46	-	80	42	-
Philadelphia	55	47	-	53	43	-
Taft	47	67	-	48	68	-
Grand Rapids	49	74	-	50	79	-
Hartford	82	71	56	70	73	48
McComb	70	47	-	69	47	-
Seattle	67	77	-	68	75	-
Portland	75	91	-	72	89	-
Jacksonville	75	75	-	75	74	-
Hammond	78	97	-	77	93	-
Bronx	63	16	-	23	12	-
			F. Grade 9			
Selmer	86	39	-	87	40	-
Athens	63	66	27	64	64	31
Wichita	62	56	29	70	56	25
Dallas	60	77	-	55	76	-
Anchorage	40	45	-	45	40	-
Rockland	69	69	81	73	67	83
Las Vegas	40	44	-	35	46	-
Fresno	63	44	-	60	48	-
Philadelphia	49	36	-	53	33	-
Taft	42	25	-	42	23	-
Grand Rapids	55	63	-	58	58	-
Hartford	70	38	-	68	34	-
McComb	71	62	-	71	60	-
Seattle	62	75	-	61	75	-
Portland	70	91	-	70	89	-
Jacksonville	77	72	20	73	73	19
Hammond	73	94	-	73	93	-
Bronx	-	-	-	-	-	-

64

TABLE N-2. NUMBER OF FULL-YEAR EXPERIMENTAL (E) AND CONTROL (C) STUDENTS GIVEN A RETENTION TEST HAVING BOTH A PRETEST AND RETENTION TEST SCORE, BY SITE, GRADE AND SUBJECT AREA

	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
Site	E	C	E	C	E	C	E	C	E	C	E	C
					A. Reading							
Selmer	86	32	80	35	83	43	-	-	-	-	-	-
Athens	46	53	-	-	61	60	57	63	-	-	-	-
Dallas	26	32	-	-	39	45	30	24	17	48	-	-
Anchorage	20	34	34	30	27	30	-	-	24	41	-	-
Las Vegas	-	-	22	19	38	17	-	-	-	-	-	-
Fresno	-	-	-	-	-	-	-	-	-	-	-	-
Grand Rapids	-	-	-	-	-	-	-	-	35	50	31	35
Portland	-	-	-	-	-	-	-	-	-	-	-	-
Jacksonville	76	48	80	39	74	51	-	-	-	-	-	-
Mesa	-	-	53	41	55	40	-	-	-	-	-	-
					B. Mathematics							
Selmer	-	-	82	34	81	41	-	-	-	-	-	-
Athens	43	56	-	-	56	58	51	62	-	-	44	48
Dallas	28	21	42	19	37	46	-	-	17	42	-	-
Anchorage	19	34	36	30	26	28	-	-	21	38	-	-
Las Vegas	-	-	23	20	34	16	-	-	-	-	-	-
Fresno	43	76	66	56	60	62	-	-	-	-	-	-
Grand Rapids	-	-	-	-	-	-	41	52	28	55	-	-
Portland	51	62	66	65	61	58	-	-	-	-	-	-
Jacksonville	76	47	79	35	76	49	-	-	-	-	-	-
Mesa	-	-	52	32	55	39	-	-	-	-	-	-

TABLE N-3. NUMBER OF EXPERIMENTAL (E) AND CONTROL (C) FULL-YEAR STUDENTS HAVING BOTH 1969-70 AND 1970-71 ATTENDANCE DATA, BY SITE AND GRADE

	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
Site	E	C	E	C	E	C	E	C	E	C	E	C
Selmer	-	-	82	29	80	36	58	25	61	39	61	28
Athens	-	-	7	37	36	24	72	66	69	68	63	64
Wichita	-	-	17	18	16	14	-	-	84	28	84	32
Dallas	-	-	69	69	71	75	65	67	70	80	63	79
Rockland	81	60	70	66	90	59	92	60	81	62	66	51
Las Vegas	-	-	-	-	-	-	31	47	44	47	25	49
Fresno	-	-	77	80	81	76	77	71	85	75	71	69
McComb	-	-	59	32	58	37	54	37	62	30	50	32
Seattle	-	-	46	64	58	63	76	65	71	68	75	61
Mesa	-	-	73	50	65	58	63	56	72	49	76	61

TABLE N-4. NUMBER OF SPECIAL TREATMENT FULL-YEAR STUDENTS HAVING BOTH A PRETEST AND POSTTEST SCORE, BY PROGRAM, SITE, GRADE, AND SUBJECT AREA

	Grade 1		Grade 2		Grade 3		Grade 7		Grade 8		Grade 9	
Program	R	M	R	M	R	M	R	M	R	M	R	M
					Grand Rapids							
Reading Centers	-	-	24	24	30	28	22	20	-	-	-	-
COMAS	-	-	-	-	-	-	110	98	35	28	27	21
Westinghouse	27	27	28	28	26	25	-	-	-	-	-	-
Project Read	26	23	41	36	54	53	-	-	-	-	-	-
					Hartford							
Waverly	39	39	60	57	58	64	-	-	-	-	-	-
Project Concern	38	37	31	17	7	8	-	-	-	-	-	-

R = Reading
M = Mathematics

APPENDIX O

MEANS AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SCORES FOR EXPERIMENTAL AND COMPARISON STUDENTS, BY SITE, GRADE, AND SUBJECT AREA

68

TABLE O-1. MEANS (AND ASSOCIATED GRADE EQUIVALENTS) AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SCORES FOR FULL-YEAR EXPERIMENTAL AND COMPARISON STUDENTS HAVING BOTH OF THESE SCORES, BY SITE, GRADE, AND SUBJECT AREA

		Experimental			Comparison		
Site		Mean	SD[b]	N[a]	Mean	SD	N
		A. GRADE 1 - READING					
Wichita	Pretest	73.24(3)[c]	12.29	71	89.39(4)	12.70	34
	Posttest	53.34(.6)	15.31		74.09(1.9)	10.42	
Rockland	Pretest	85.27(4)	14.40	82	105.83(6)	20.30	76
	Posttest	68.57(1.6)	12.96		89.09(2.6)	17.15	
McComb	Pretest	50.09(1)	17.89	35	81.77(3)	16.83	31
	Posttest	56.09(.8)	10.00		66.26(1.5)	9.97	
Jacksonville	Pretest	64.68(2)	19.48	90	59.62(2)	15.58	45
	Posttest	58.76(1.0)	16.04		57.29(.8)	16.46	
Hammond	Pretest	74.69(3)	18.16	74	82.88(4)	22.98	34
	Posttest	60.32(1.0)	17.03		62.88(1.2)	20.83	
		B. GRADE 1 - MATHEMATICS					
Wichita	Pretest	73.24(3)	12.29	71	88.79(4)	12.70	34
	Posttest	36.39(1.1)	10.49		50.62(1.9)	10.58	
Rockland	Pretest	85.26(4)	14.45	83	105.83(6)	20.30	76
	Posttest	49.58(1.8)	13.20		66.13(2.6)	15.91	
McComb	Pretest	50.09(1)	17.89	35	81.77(3)	16.83	31
	Posttest	41.14(1.4)	9.89		41.13(1.4)	14.62	
Jacksonville	Pretest	65.08(2)	19.21	89	59.62(2)	15.58	45
	Posttest	39.97(1.3)	12.80		33.80(1.0)	12.34	
Hammond	Pretest	74.25(3)	18.42	75	82.88(4)	22.98	34
	Posttest	43.12(1.5)	14.51		50.15(1.8)	20.09	

(a) N = Number of students.

(b) SD - Standard deviation.

(c) Stanine values are given for Grade 1 pretest means.

TABLE O-1. (Continued)

		C. GRADE 2 - READING					
		Experimental			Comparison		
Site		Mean	SD[b]	N[a]	Mean	SD	N
Wichita	Pretest	32.18(1.5)	10.48	65	43.04(1.8)	16.40	44
	Posttest	57.48(2.0)	13.71		63.66(2.2)	13.45	
Rockland	Pretest	38.06(1.7)	10.92	80	67.38(2.3)	9.56	63
	Posttest	60.61(2.1)	14.84		74.59(3.4)	2.03	
McComb	Pretest	24.90(1.3)	10.72	48	45.82(1.9)	17.01	27
	Posttest	44.29(1.8)	14.81		63.15(2.2)	16.41	
Jacksonville	Pretest	29.35(1.4)	11.89	93	32.06(1.5)	14.36	48
	Posttest	51.84(1.9)	17.78		51.29(1.9)	19.18	
Hammond	Pretest	31.04(1.5)	11.44	75	40.54(1.8)	15.28	59
	Posttest	50.63(1.9)	16.26		59.25(2.1)	16.16	
		D. GRADE 2 - MATHEMATICS					
Wichita	Pretest	27.48(1.3)	11.56	65	34.87(1.5)	12.75	38
	Posttest	39.82(1.8)	12.41		46.76(2.2)	9.52	
Rockland	Pretest	33.14(1.5)	10.82	79	46.14(2.2)	6.55	62
	Posttest	49.63(2.4)	9.06		59.36(3.8)	3.10	
McComb	Pretest	19.12(1.2)	12.87	49	38.08(1.7)	14.42	26
	Posttest	37.02(1.6)	9.97		51.15(2.4)	10.31	
Jacksonville	Pretest	22.60(1.2)	8.33	88	25.48(1.3)	9.44	44
	Posttest	43.25(2.0)	10.73		44.43(2.0)	10.57	
Hammond	Pretest	29.73(1.4)	10.93	73	37.59(1.7)	12.62	58
	Posttest	43.45(2.0)	10.21		47.53(2.3)	9.72	

TABLE 0-1. (Continued)

Site		Experimental Mean	Experimental SD$^{(b)}$	Experimental N$^{(a)}$	Comparison Mean	Comparison SD	Comparison N
		E. GRADE 3 - READING					
Wichita	Pretest	32.02(2.1)	10.55	63	48.18(2.4)	21.23	38
	Posttest	43.71(2.4)	13.95		60.87(2.9)	18.16	
Rockland	Pretest	34.99(2.2)	9.63	92	70.53(3.1)	9.40	108
	Posttest	58.89(2.8)	16.11		78.23(4.1)	5.37	
McComb	Pretest	25.54(1.8)	7.93	69	44.27(2.4)	22.23	37
	Posttest	40.12(2.3)	15.04		56.68(2.8)	20.22	
Jacksonville	Pretest	30.36(2.0)	12.87	91	38.49(2.2)	16.89	47
	Posttest	44.40(2.4)	17.14		51.74(2.6)	18.79	
Hammond	Pretest	35.37(2.2)	16.26	74	52.10(2.5)	20.39	51
	Posttest	48.62(2.5)	18.10		62.10(3.0)	18.86	
		F. GRADE 3 - MATHEMATICS					
Wichita	Pretest	41.38(2.0)	14.95	66	53.66(2.4)	22.84	35
	Posttest	54.64(2.4)	15.81		70.94(3.0)	21.21	
Rockland	Pretest	50.13(2.3)	15.85	89	73.92(3.0)	14.10	110
	Posttest	73.27(3.1)	17.63		93.40(4.2)	10.02	
McComb	Pretest	35.46(1.6)	10.00	69	52.22(2.3)	18.77	37
	Posttest	53.35(2.3)	14.36		69.22(3.0)	21.26	
Jacksonville	Pretest	38.04(1.8)	12.49	92	41.11(2.0)	15.90	45
	Posttest	58.84(2.5)	19.10		60.49(2.5)	20.05	
Hammond	Pretest	46.68(2.2)	15.02	73	56.71(2.4)	17.02	49
	Posttest	60.88(2.6)	19.02		73.49(3.1)	18.67	

TABLE 0-1. (Continued)

Site		Experimental Mean	Experimental SD(b)	Experimental N(a)	Comparison Mean	Comparison SD	Comparison N
		G. GRADE 7 - READING					
Athens	Pretest	34.73(4.2)	13.59	64	64.47(6.6)	15.75	38
	Posttest	39.98(4.5)	18.28		70.42(7.1)	13.79	
Wichita	Pretest	42.57(4.8)	13.80	72	53.30(5.6)	11.95	27
	Posttest	47.00(5.0)	16.22		55.00(5.7)	12.67	
Rockland	Pretest	48.21(5.2)	16.05	80	77.69(8.2)	10.97	101
	Posttest	54.12(5.6)	15.58		82.40(9.3)	9.12	
Hartford	Pretest	37.94(4.4)	13.78	71	29.82(3.7)	15.54	45
	Posttest	44.97(4.9)	15.13		36.29(4.3)	16.36	
Jacksonville	Pretest	31.03(3.8)	11.54	75	30.80(3.8)	11.52	15
	Posttest	33.65(4.1)	14.30		35.60(4.3)	10.38	
		H. GRADE 7 - MATHEMATICS					
Athens	Pretest	33.95(4.2)	10.52	61	63.06(5.9)	15.59	35
	Posttest	49.31(5.1)	18.22		80.14(6.9)	19.99	
Wichita	Pretest	41.07(4.6)	14.90	74	54.00(5.3)	13.45	29
	Posttest	48.55(5.1)	16.62		61.48(5.7)	18.01	
Rockland	Pretest	54.07(5.3)	17.49	82	83.90(7.4)	15.20	99
	Posttest	65.61(6.0)	18.06		97.31(9.0)	12.21	
Hartford	Pretest	40.81(4.6)	14.73	62	34.03(4.2)	15.93	32
	Posttest	52.39(5.2)	16.45		49.50(5.1)	20.15	
Jacksonville	Pretest	35.38(4.2)	13.60	74	35.06(4.2)	14.69	17
	Posttest	43.91(4.7)	15.41		39.65(4.5)	13.82	

TABLE 0-1. (Continued)

I. GRADE 8 - READING

Site		Experimental			Comparison		
		Mean	SD(b)	N(a)	Mean	SD	N
Athens	Pretest	24.84(3.9)	6.58	51	46.50(6.7)	11.19	26
	Posttest	28.57(4.4)	8.96		56.19(7.6)	13.07	
Wichita	Pretest	35.95(5.3)	12.86	65	45.87(6.6)	12.10	23
	Posttest	42.43(6.2)	13.98		51.52(7.3)	10.11	
Rockland	Pretest	37.75(5.6)	12.19	73	70.20(9.8)	15.79	103
	Posttest	46.49(6.6)	13.94		76.36(11.6)	11.95	
Hartford	Pretest	33.13(4.9)	14.24	82	28.23(4.3)	15.92	56
	Posttest	41.01(6.0)	14.26		35.77(5.3)	15.47	

J. GRADE 8 - MATHEMATICS

Site		Experimental			Comparison		
		Mean	SD(b)	N(a)	Mean	SD	N
Athens	Pretest	31.42(4.5)	8.99	45	48.93(6.2)	10.30	27
	Posttest	34.53(5.2)	11.12		67.59(8.0)	11.82	
Wichita	Pretest	35.39(4.9)	10.00	69	45.72(6.0)	12.32	18
	Posttest	44.30(6.0)	11.75		56.78(7.1)	15.69	
Rockland	Pretest	46.61(6.0)	14.94	71	81.00(8.8)	16.31	104
	Posttest	56.25(6.9)	17.84		90.88(11.0)	15.60	
Hartford	Pretest	41.24(5.6)	13.83	70	32.81(4.7)	15.15	48
	Posttest	45.07(6.2)	14.40		40.06(5.7)	15.88	

TABLE 0-1. (Continued)

Site		Experimental Mean	Experimental SD[b]	Experimental N[a]	Comparison Mean	Comparison SD	Comparison N
K. GRADE 9 - READING							
Athens	Pretest	30.30(4.5)	9.64	63	60.22(8.2)	13.40	27
	Posttest	37.87(5.6)	15.15		66.41(9.3)	12.69	
Wichita	Pretest	44.76(6.4)	14.27	62	52.00(7.3)	15.81	29
	Posttest	49.48(6.9)	15.55		56.52(7.8)	14.73	
Rockland	Pretest	48.75(6.9)	17.47	69	76.28(11.6)	12.70	81
	Posttest	57.75(8.0)	16.70		78.84(12.5)	12.53	
Jacksonville	Pretest	33.00(4.9)	12.61	77	33.65(5.0)	13.10	20
	Posttest	41.19(6.0)	14.35		40.75(6.0)	12.12	
L. GRADE 9 - MATHEMATICS							
Athens	Pretest	37.41(5.2)	10.91	64	72.29(8.0)	17.99	31
	Posttest	49.30(6.6)	18.68		85.48(10.1)	15.80	
Wichita	Pretest	43.53(5.9)	13.84	70	53.88(6.7)	16.04	25
	Posttest	51.80(6.7)	17.18		56.88(7.1)	16.79	
Rockland	Pretest	55.48(6.7)	19.90	73	90.31(10.4)	16.49	83
	Posttest	65.26(7.8)	20.49		95.39(12.2)	15.74	
Jacksonville	Pretest	41.52(5.7)	13.61	73	47.26(6.0)	16.39	19
	Posttest	48.32(6.4)	15.20		50.58(6.7)	18.12	

APPENDIX P

MEANS AND STANDARD DEVIATIONS OF PRETEST AND SECOND-YEAR TEST SCORES FOR FULL-YEAR EXPERIMENTAL AND CONTROL STUDENTS HAVING BOTH OF THESE SCORES, BY SITE, GRADE, SUBJECT, AND GROUP

75

TABLE P-1. MEANS (AND ASSOCIATED GRADE EQUIVALENTS) AND STANDARD DEVIATIONS OF PRETEST AND SECOND YEAR TEST SCORES FOR FULL-YEAR EXPERIMENTAL AND CONTROL STUDENTS HAVING BOTH OF THESE SCORES, BY SITE, GRADE, SUBJECT, AND GROUP

		A. GRADE 1 - READING					
		Experimental			Control		
Site		Mean	SD[b]	N[a]	Mean	SD	N
Selmer	Pretest	69.95(2)[c]	16.99	86	74.22(3)	22.07	32
	2nd Yr. Test	80.08(2.0)	16.67		73.78(1.7)	18.44	
Athens	Pretest	58.15(2)	18.17	46	72.59(3)	19.70	53
	2nd Yr. Test	58.65(0.7)	10.64		67.87(1.3)	15.90	
Dallas	Pretest	46.96(1)	19.77	26	45.31(1)	19.63	32
	2nd Yr. Test	63.27(1.0)	15.67		72.37(1.6)	17.46	
Anchorage	Pretest	75.00(3)	11.783	20	93.41(5)	15.09	34
	2nd Yr. Test	73.30(1.7)	14.568		85.35(2.2)	13.81	
Jacksonville	Pretest	63.66(2)	19.82	76	56.42(2)	17.17	48
	2nd Yr. Test	67.97(1.3)	18.47		55.29(<.6)	11.54	
		B. GRADE 1 - MATHEMATICS					
Athens	Pretest	59.81(2)	17.51	43	73.02(3)	19.66	56
	2nd Yr. Test	44.02(1.3)	12.98		56.18(1.8)	16.72	
Dallas	Pretest	45.43(1)	19.98	28	46.71(1)	19.33	21
	2nd Yr. Test	48.57(1.6)	18.35		54.95(1.7)	17.90	
Anchorage	Pretest	74.79(3)	12.07	19	93.41(5)	15.09	34
	2nd Yr. Test	49.16(1.6)	14.42		65.41(2.2)	13.10	
Fresno	Pretest	85.77(4)	16.66	43	81.93(3)	16.34	76
	2nd Yr. Test	48.88(1.6)	12.28		52.03(1.7)	15.38	
Portland	Pretest	80.10(3)	20.94	51	88.63(4)	14.74	62
	2nd Yr. Test	53.53(1.7)	21.12		64.63(2.1)	14.89	
Jacksonville	Pretest	63.66(2)	19.82	76	56.66(2)	16.29	47
	2nd Yr. Test	48.07(1.7)	17.08		32.26(0.7)	11.81	

(a) N = Number of subjects.
(b) SD = Standard deviation.
(c) Stanine values are given for Grade 1 pretest means throughout.

TABLE P-1. (Continued)

		Experimental			Control		
Site		Mean	SD(b)	N(a)	Mean	SD	N
		C. GRADE 2 - READING					
Selmer	Pretest	42.09(1.8)	15.72	80	33.80(1.4)	10.54	35
	2nd Yr. Test	70.23(2.4)	9.95		63.14(2.2)	13.61	
Anchorage	Pretest	31.15(1.5)	7.95	34	42.33(1.8)	13.55	30
	2nd Yr. Test	64.35(2.2)	12.22		69.07(2.4)	9.03	
Las Vegas	Pretest	24.00(1.3)	7.96	22	23.58(1.3)	6.65	19
	2nd Yr. Test	44.86(1.9)	13.50		53.47(2.0)	16.65	
Jacksonville	Pretest	29.95(1.4)	12.50	80	25.59(1.3)	10.36	39
	2nd Yr. Test	60.06(2.2)	15.61		52.03(2.0)	14.81	
Mesa	Pretest	32.83(1.5)	10.66	53	29.73(1.4)	13.61	41
	2nd Yr. Test	56.90(2.1)	15.66		51.15(2.0)	17.37	
		D. GRADE 2 - MATHEMATICS					
Selmer	Pretest	34.85(1.5)	11.05	82	37.32(1.6)	10.27	34
	2nd Yr. Test	54.82(3.0)	9.30		54.59(3.0)	8.58	
Dallas	Pretest	23.38(1.2)	12.24	42	20.47(1.2)	6.91	19
	2nd Yr. Test	44.12(2.0)	12.16		30.00(1.4)	11.88	
Anchorage	Pretest	26.92(1.3)	8.08	36	38.20(1.7)	9.31	30
	2nd Yr. Test	52.50(2.6)	7.55		55.60(3.0)	5.73	
Las Vegas	Pretest	18.70(1.2)	9.00	23	22.85(1.2)	8.02	20
	2nd Yr. Test	35.74(1.6)	10.93		46.40(2.2)	9.19	
Fresno	Pretest	30.23(1.4)	10.23	66	33.55(1.5)	10.17	56
	2nd Yr. Test	48.33(2.3)	9.09		53.86(2.7)	6.01	
Portland	Pretest	39.18(1.7)	11.04	66	40.08( )	8.46	65
	2nd Yr. Test	53.62(2.7)	6.78		55.94(3.1)	5.35	
Jacksonville	Pretest	22.67(1.2)	8.53	79	21.23(1.2)	6.28	35
	2nd Yr. Test	46.04(2.2)	10.60		39.71(1.8)	9.14	
Mesa	Pretest	31.08(1.4)	8.94	52	32.47(1.5)	10.83	32
	2nd Yr. Test	43.14(2.0)	10.74		44.44(2.0)	7.73	

E. GRADE 3 - READING

Site		Experimental			Control		
		Mean	SD(b)	N(a)	Mean	SD	N
Selmer	Pretest	43.96(2.4)	16.63	83	46.72(2.4)	17.68	43
	2nd Yr. Test	71.61(3.2)	12.02		66.93(2.9)	15.55	
Athens	Pretest	36.54(2.2)	15.58	61	42.43(2.3)	18.02	60
	2nd Yr. Test	59.02(2.7)	18.26		58.75(2.7)	19.34	
Dallas	Pretest	30.21(2.0)	10.89	39	25.04(1.7)	10.29	45
	2nd Yr. Test	43.15(2.4)	17.53		32.18(2.1)	15.99	
Anchorage	Pretest	36.89(2.2)	9.59	27	49.27(2.4)	17.76	30
	2nd Yr. Test	63.07(2.8)	12.61		69.03(3.0)	13.36	
Las Vegas	Pretest	28.42(1.9)	9.82	38	30.29(2.0)	8.48	17
	2nd Yr. Test	46.29(2.4)	16.83		58.18(2.6)	13.36	
Jacksonville	Pretest	30.80(2.0)	12.14	74	29.02(1.9)	10.63	51
	2nd Yr. Test	49.49(2.4)	20.00		42.82(2.4)	18.08	
Mesa	Pretest	40.75(2.3)	15.24	55	37.60(2.2)	13.64	40
	2nd Yr. Test	58.73(2.7)	18.17		58.85(2.7)	17.156	

F. GRADE 3 - MATHEMATICS

Site		Experimental			Control		
		Mean	SD(b)	N(a)	Mean	SD	N
Selmer	Pretest	55.00(2.4)	16.42	81	64.76(2.6)	14.52	41
	2nd Yr. Test	88.26(3.7)	11.90		82.51(3.3)	14.50	
Athens	Pretest	46.48(2.2)	15.50	56	55.53(2.4)	16.55	58
	2nd Yr. Test	74.86(3.0)	15.66		77.05(3.1)	20.14	
Dallas	Pretest	35.97(1.7)	12.65	37	35.48(1.6)	10.59	46
	2nd Yr. Test	57.70(2.4)	19.42		48.96(2.3)	17.54	
Anchorage	Pretest	47.92(2.3)	9.00	26	57.04(2.4)	15.62	28
	2nd Yr. Test	70.39(2.8)	14.76		79.71(3.2)	14.23	
Las Vegas	Pretest	35.27(1.6)	9.94	34	40.00(1.9)	11.48	16
	2nd Yr. Test	59.59(2.5)	15.19		70.88(2.9)	16.96	
Fresno	Pretest	47.37(2.2)	11.76	60	54.08(2.4)	15.06	62
	2nd Yr. Test	73.45(3.0)	15.57		78.81(3.2)	14.23	
Portland	Pretest	51.93(2.3)	12.61	61	64.14(2.6)	15.44	58
	2nd Yr. Test	76.38(3.0)	15.10		88.57(3.8)	12.83	
Jacksonville	Pretest	37.74(1.8)	11.86	76	36.39(1.7)	10.13	49
	2nd Yr. Test	64.47(2.6)	17.26		54.25(2.4)	15.74	
Mesa	Pretest	49.49(2.3)	15.54	55	45.36(2.2)	12.07	39
	2nd Yr. Test	65.07(2.6)	17.62		63.03(2.6)	17.97	

TABLE P-1. (Continued)

G. GRADE 7 - READING

		Experimental			Control		
Site		Mean	SD(b)	N(a)	Mean	SD	N
Athens	Pretest	34.93(4.2)	11.97	57	42.38(4.8)	15.78	63
	2nd Yr. Test	41.28(4.6)	16.70		44.41(4.9)	17.48	
Dallas	Pretest	26.60(3.4)	9.93	30	28.71(3.6)	9.01	24
	2nd Yr. Test	31.00(3.8)	13.53		32.54(3.9)	13.78	

H. GRADE 7 - MATHEMATICS

		Experimental			Control		
Site		Mean	SD(b)	N(a)	Mean	SD	N
Athens	Pretest	35.39(4.2)	10.44	51	45.86(4.8)	16.25	62
	2nd Yr. Test	48.10(4.9)	17.66		52.10(5.2)	19.48	
Grand Rapids	Pretest	38.39(4.4)	11.751	41	39.85(4.5)	16.08	52
	2nd Yr. Test	51.49(5.2)	15.30		56.81(5.4)	18.52	

TABLE P-1. (Continued)

I. GRADE 8 - READING

Site		Experimental			Control		
		Mean	SD(b)	N(a)	Mean	SD	N
Dallas	Pretest	17.53(3.0)	6.24	17	24.79(3.9)	7.61	43
	2nd Yr. Test	27.59(4.3)	8.19		29.54(4.5)	12.86	
Anchorage	Pretest	38.29(5.6)	15.09	24	51.12(7.1)	15.53	41
	2nd Yr. Test	39.79(5.8)	13.06		53.17(7.3)	15.72	
Grand Rapids	Pretest	27.89(4.3)	10.11	35	34.38(5.0)	10.99	50
	2nd Yr. Test	39.66(5.8)	12.98		39.38(5.7)	13.17	

J. GRADE 8 - MATHEMATICS

Site		Experimental			Control		
		Mean	SD(b)	N(a)	Mean	SD	N
Dallas	Pretest	22.06(3.8)	7.65	17	29.00(4.3)	8.50	42
	2nd Yr. Test	37.82(5.2)	11.66		41.41(5.6)	15.78	
Anchorage	Pretest	39.48(5.4)	15.69	21	48.84(6.2)	11.92	38
	2nd Yr. Test	49.14(6.2)	20.09		52.29(6.6)	19.07	
Grand Rapids	Pretest	37.93(5.3)	11.08	28	36.60(5.2)	10.54	55
	2nd Yr. Test	51.21(6.4)	16.26		52.91(6.6)	18.07	

TABLE P-1. (Continued)

Site		Experimental Mean	Experimental SD[b]	Experimental N[a]	Control Mean	Control SD	Control N
		K. GRADE 9 - READING					
Grand Rapids	Pretest	37.42(5.5)	10.25	31	39.17(5.7)	10.34	35
	2nd Yr. Test	49.48(6.9)	12.46		46.00(6.6)	13.17	
		L. GRADE 9 - MATHEMATICS					
Athens	Pretest	36.07(5.1)	10.23	44	42.56(5.7)	20.014	48
	2nd Yr. Test	51.27(6.4)	17.80		54.54(6.7)	23.05	

APPENDIX Q

MEANS AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SCORES FOR FULL-YEAR EXPERIMENTAL AND SPECIAL TREATMENT STUDENTS, BY PROGRAM, SITE, GRADE, AND SUBJECT AREA

TABLE Q-1. MEANS (AND ASSOCIATED GRADE EQUIVALENTS) AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SCORES FOR FULL-YEAR EXPERIMENTAL AND SPECIAL TREATMENT STUDENTS HAVING BOTH OF THESE SCORES, BY PROGRAM, SITE, GRADE, AND SUBJECT AREA

		Experimental			Special Treatment		
Program		Mean	SD[b]	N[a]	Mean	SD	N
		A. GRAND RAPIDS - GRADE 1 - READING					
Westinghouse	Pretest	59.00(2)[c]	21.86	60	74.11(3)	12.84	27
	Posttest	56.14(.8)	16.45		60.26(1.0)	12.16	
Project Read	Pretest	59.00(2)	21.86	60	71.04(3)	16.72	26
	Posttest	56.14(.8)	16.45		66.23(1.5)	20.51	
		B. HARTFORD - GRADE 1 - READING					
Waverly	Pretest	66.17(2)	20.40	40	70.67(3)	23.69	39
	Posttest	57.46(.8)	14.65		66.10(1.5)	12.77	
Project Concern	Pretest	66.17(2)	20.40	40	75.97(3)	16.62	38
	Posttest	57.46(.8)	14.65		55.89(.8)	24.80	
		C. GRAND RAPIDS - GRADE 1 - MATHEMATICS					
Westinghouse	Pretest	61.73(2)	22.31	60	74.11(3)	12.84	27
	Posttest	30.54(.8)	10.31		36.26(1.1)	10.23	
Project Read	Pretest	61.73(2)	22.31	60	71.26(3)	17.49	23
	Posttest	30.54(.8)	10.31		42.83(1.5)	18.03	
		D. HARTFORD - GRADE 1 - MATHEMATICS					
Waverly	Pretest	63.11(2)	20.34	40	71.26(3)	23.41	39
	Posttest	31.31(.8)	13.54		45.21(1.6)	12.91	
Project Concern	Pretest	63.11(2)	20.34	40	75.92(3)	16.85	37
	Posttest	31.31(.8)	13.54		35.86(1.1)	13.65	

(a) N = Number of students.

(b) SD = Standard deviation

(c) Stanine values are given for Grade 1 pretest means.

TABLE Q-1. (Continued)

		Experimental			Special Treatment		
Program		Mean	SD[b]	N[a]	Mean	SD	N
		E. GRAND RAPIDS - GRADE 2 - READING					
Reading Centers	Pretest	24.14(1.3)	8.97	71	32.50(1.5)	6.37	24
	Posttest	44.11(1.8)	15.72		61.96(2.2)	11.31	
Westinghouse	Pretest	24.14(1.3)	8.97	71	28.21(1.4)	10.75	28
	Posttest	44.11(1.8)	15.72		48.71(1.9)	16.15	
Project Read	Pretest	24.14(1.3)	8.97	71	34.22(1.5)	15.87	41
	Posttest	44.11(1.8)	15.72		54.98(2.0)	17.54	
		F. HARTFORD - GRADE 2 - READING					
Waverly	Pretest	22.54(1.3)	8.70	54	29.73(1.4)	16.12	60
	Posttest	39.81(1.7)	13.65		54.13(2.0)	16.43	
Project Concern	Pretest	22.54(1.3)	8.70	54	39.77(1.7)	20.49	31
	Posttest	39.81(1.7)	13.65		58.61(2.1)	14.12	
		G. GRAND RAPIDS - GRADE 2 - MATHEMATICS					
Reading Centers	Pretest	33.45(1.5)	9.99	53	29.08(1.4)	9.84	24
	Posttest	44.57(2.1)	9.72		46.79(2.2)	9.13	
Westinghouse	Pretest	33.45(1.5)	9.99	53	27.00(1.3)	11.18	28
	Posttest	44.57(2.1)	9.72		40.89(1.9)	9.52	
Project Read	Pretest	33.45(1.5)	9.99	53	30.64(1.4)	12.66	36
	Posttest	44.57(2.1)	9.72		45.06(2.1)	11.34	
		H. HARTFORD - GRADE 2 - MATHEMATICS					
Waverly	Pretest	18.96(1.2)	7.74	49	24.58(1.3)	12.17	57
	Posttest	34.39(1.5)	9.12		47.44(2.2)	9.18	
Project Concern	Pretest	18.96(1.2)	7.74	49	37.00(1.6)	12.55	17
	Posttest	34.39(1.5)	9.12		46.47(2.2)	10.72	

TABLE Q-1. (Continued)

Program		Experimental Mean	Experimental SD$^{(b)}$	Experimental N$^{(a)}$	Special Treatment Mean	Special Treatment SD	Special Treatment N
		I. GRAND RAPIDS - GRADE 3 - READING					
Reading Centers	Pretest	33.12(2.1)	12.73	81	26.27(1.8)	6.69	30
	Posttest	44.72(2.4)	17.27		50.77(2.6)	13.44	
Westinghouse	Pretest	33.12(2.1)	12.73	81	35.23(2.2)	16.61	26
	Posttest	44.72(2.4)	17.27		52.15(2.6)	16.07	
Project Read	Pretest	33.12(2.1)	12.73	81	23.98(1.7)	6.00	54
	Posttest	44.72(2.4)	17.27		42.87(2.4)	12.78	
		J. HARTFORD - GRADE 3 - READING					
Waverly	Pretest	24.96(1.7)	9.59	53	32.09(2.1)	18.02	58
	Posttest	44.89(2.4)	15.06		53.40(2.6)	19.12	
Project Concern	Pretest	24.96(1.7)	9.59	53	56.86(2.6)	27.49	7
	Posttest	44.89(2.4)	15.06		62.86(3.0)	13.63	
		K. GRAND RAPIDS - GRADE 3 - MATHEMATICS					
Reading Centers	Pretest	40.64(2.0)	13.45	78	39.21(1.9)	13.82	28
	Posttest	57.62(2.4)	15.97		64.04(2.7)	18.12	
Westinghouse	Pretest	40.64(2.0)	13.45	78	41.56(2.0)	14.43	25
	Posttest	57.62(2.4)	15.97		58.80(2.5)	16.26	
Project Read	Pretest	40.64(2.0)	13.45	78	34.79(1.6)	10.79	53
	Posttest	57.62(2.4)	15.97		65.11(2.7)	13.44	
		L. HARTFORD - GRADE 3 - MATHEMATICS					
Waverly	Pretest	41.02(2.0)	14.51	48	38.42(1.8)	18.58	64
	Posttest	48.92(2.3)	13.64		68.23(2.9)	19.34	
Project Concern	Pretest	41.02(2.0)	14.51	48	43.88(2.2)	8.98	8
	Posttest	48.92(2.3)	13.64		60.13(2.5)	11.46	

TABLE Q-1. (Continued)

Program		Experimental			Special Treatment		
		Mean	SD(b)	N(a)	Mean	SD	N
		M. GRAND RAPIDS - GRADE 7 — READING					
Reading Centers	Pretest	32.45(3.8)	11.62	59	37.09(4.3)	14.32	22
	Posttest	37.85(4.4)	12.79		38.86(4.5)	15.21	
COMAS	Pretest	32.45(3.8)	11.62	59	33.56(4.1)	11.20	110
	Posttest	37.85(4.4)	12.79		41.98(4.8)	14.69	
		N. GRAND RAPIDS - GRADE 7 - MATHEMATICS					
Reading Centers	Pretest	36.46(4.3)	13.50	60	41.35(4.6)	18.16	20
	Posttest	43.77(4.7)	14.32		48.80(5.1)	19.12	
COMAS	Pretest	36.46(4.3)	13.50	60	32.14(4.0)	9.40	98
	Posttest	43.77(4.7)	14.32		43.69(4.7)	13.14	
		O. GRAND RAPIDS - GRADE 8 - READING					
COMAS	Pretest	26.82(4.2)	10.04	49	23.09(3.7)	9.98	35
	Posttest	35.77(5.3)	12.10		31.71(4.8)	11.72	
		P. GRAND RAPIDS - GRADE 8 - MATHEMATICS					
COMAS	Pretest	34.68(4.9)	11.45	50	31.11(4.5)	7.40	28
	Posttest	37.31(5.4)	12.16		33.14(4.9)	9.02	
		Q. GRAND RAPIDS - GRADE 9 - READING					
COMAS	Pretest	33.97(5.0)	10.37	55	26.07(4.1)	6.44	27
	Posttest	40.63(6.0)	13.72		30.96(4.6)	8.29	
		R. GRAND RAPIDS - GRADE 9 - MATHEMATICS					
COMAS	Pretest	41.31(5.6)	12.89	58	30.10(4.4)	10.05	21
	Posttest	45.73(6.2)	15.22		34.10(5.1)	10.40	

APPENDIX R

ESTIMATED POSTTEST MEANS FOR EXPERIMENTAL (E) AND CONTROL (C) GROUPS, DIFFERENCES BETWEEN THESE MEANS, AND ASSOCIATED t RATIOS, FOR ALL SITES, GRADES, AND SUBJECTS

TABLE R-1 . PRE-POST MODEL RESULTS

GRADE 1-READING

Site	Mean Pretest Value*	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	(3)	68 (1.6)	55 (0.7)	+13 (0.9)	5.857
Athens	(2)	53 (0.6)	54 (0.7)	-1 (-0.1)	-0.424
Wichita	(3)	53 (0.6)	69 (1.6)	-16 (-1.0)	-5.663
Dallas	(1)	53 (0.6)	49 (<0.6)	+4 (>0)	1.397
Anchorage	(4)	62 (1.2)	65 (1.4)	-3 (-0.2)	-0.729
Rockland	(4)	71 (1.7)	74 (1.9)	-3 (-0.2)	-1.608
Las Vegas	(2)	51 (<0.6)	57 (0.8)	-6 (<-0.2)	-2.120
Fresno	(3)	57 (0.8)	56 (0.8)	+1 (0)	0.470
Philadelphia	(2)	51 (<0.6)	57 (0.8)	-6 (<-0.2)	-2.667
Taft	(2)	62 (1.2)	62 (1.2)	0 (0)	0.135
Grand Rapids	(2)	53 (0.6)	58 (0.9)	-5 (-0.3)	-1.701
Hartford	(2)	50 (<0.6)	55 (0.7)	-5 (<-0.1)	-1.760
McComb	(2)	60 (1.0)	62 (1.2)	-2 (-0.2)	-0.573
Seattle	(4)	63 (1.2)	73 (1.8)	-10 (-0.6)	-3.250
Portland	(4)	62 (1.2)	71 (1.7)	-9 (-0.5)	-4.049
Jacksonville	(2)	57 (0.8)	36 (<0.6)	+21 (>+0.2)	11.002
Hammond	(3)	62 (1.2)	73 (1.8)	-11 (-0.6)	-4.759
Bronx	---	----	----	----	----

* The stanine value corresponding to the combined E and C pretest raw score mean.

TABLE R-1. (Continued)

GRADE 1-MATHEMATICS

Site	Mean Pretest Value*	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	(3)	46 (1.7)	42 (1.5)	+4 (0.2)	1.650
Athens	(2)	40 (1.3)	39 (1.3)	+1 (0)	0.492
Wichita	(3)	36 (1.1)	39 (1.3)	-3 (-0.2)	-1.435
Dallas	(1)	34 (1.0)	29 (0.6)	+5 (0.4)	2.253
Anchorage	(4)	41 (1.4)	42 (1.5)	-1 (-0.1)	-0.219
Rockland	(4)	51 (1.9)	53 (1.9)	-2 (0)	-0.850
Las Vegas	(2)	27 (<0.6)	40 (1.3)	-13 (<-0.7)	-5.348
Fresno	(4)	33 (0.9)	42 (1.5)	-9 (-0.6)	-4.336
Philadelphia	(2)	28 (<0.6)	38 (1.2)	-10 (<-0.6)	-5.443
Taft	(2)	40 (1.3)	44 (1.6)	-4 (-0.3)	-1.547
Grand Rapids	(2)	32 (0.9)	41 (1.4)	-9 (-0.5)	-4.073
Hartford	(2)	30 (0.7)	39 (1.3)	-9 (-0.6)	-3.874
McComb	(2)	44 (1.6)	36 (1.1)	+8 (0.5)	2.616
Seattle	(4)	41 (1.4)	46 (1.7)	-5 (-0.3)	-1.830
Portland	(4)	41 (1.4)	55 (2.0)	-14 (-0.6)	-6.575
Jacksonville	(2)	38 (1.2)	24 (<0.6)	+14 (>+0.6)	9.016
Hammond	(3)	45 (1.6)	47 (1.7)	-2 (-0.1)	-0.934
Bronx	---	----	----	----	----

* The stanine value corresponding to the combined E and C pretest raw score mean.

TABLE R-1. (Continued)

GRADE 2-READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	40 (1.7)	65 (2.2)	60 (2.1)	+5 (+.1)	2.114
Athens	33 (1.5)	52 (1.9)	51 (1.9)	+1 (0)	0.405
Wichita	39 (1.7)	63 (2.2)	58 (2.0)	+5 (+.2)	2.115
Dallas	25 (1.3)	37 (1.6)	33 (1.5)	+4 (+.1)	1.579
Anchorage	36 (1.6)	63 (2.2)	58 (2.0)	+5 (+.2)	1.390
Rockland	45 (1.9)	66 (2.2)	69 (2.4)	-3 (-.2)	-1.379
Las Vegas	23 (1.3)	32 (1.5)	46 (1.8)	-14 (-.3)	-5.134
Fresno	35 (1.6)	54 (2.0)	51 (1.9)	+3 (+.1)	1.355
Philadelphia	24 (1.3)	40 (1.7)	42 (1.7)	-2 (0)	-0.791
Taft	37 (1.6)	52 (1.9)	57 (2.0)	-5 (-.1)	-2.291
Grand Rapids	30 (1.4)	50 (1.9)	53 (2.0)	-3 (-.1)	-1.380
Hartford	27 (1.4)	42 (1.7)	51 (1.9)	-9 (-.2)	-3.308
McComb	29 (1.4)	47 (1.8)	55 (2.0)	-8 (-.2)	-2.468
Seattle	39 (1.7)	58 (2.0)	65 (2.2)	-7 (-.2)	-3.048
Portland	42 (1.8)	59 (2.1)	65 (2.2)	-6 (-.1)	-3.577
Jacksonville	27 (1.4)	50 (1.9)	44 (1.8)	+6 (+.1)	2.427
Hammond	34 (1.5)	53 (2.0)	55 (2.0)	-2 (0)	-0.907
Bronx	28 (1.4)	41 (1.7)	54 (2.0)	-13 (-.3)	-4.065

TABLE R-1. (Continued)

GRADE 2-MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	35 (1.5)	49 (2.3)	47 (2.2)	+2 (+.1)	1.472
Athens	30 (1.4)	44 (2.0)	44 (2.0)	+0 (0)	-0.189
Wichita	30 (1.4)	41 (1.9)	42 (1.9)	-1 (0)	-0.600
Dallas	21 (1.2)	34 (1.5)	25 (1.3)	+9 (+.2)	4.298
Anchorage	32 (1.4)	49 (2.3)	48 (2.3)	+1 (0)	1.042
Rockland	38 (1.7)	53 (2.6)	56 (3.1)	-3 (-.5)	-2.926
Las Vegas	20 (1.2)	29 (1.4)	36 (1.6)	-7 (-.2)	-3.101
Fresno	31 (1.4)	43 (2.0)	47 (2.2)	-4 (-.2)	-2.972
Philadelphia	23 (1.3)	36 (1.6)	40 (1.8)	-4 (-.2)	-2.717
Taft	28 (1.4)	43 (2.0)	44 (2.0)	-1 (0)	-0.779
Grand Rapids	32 (1.5)	44 (2.0)	45 (2.1)	-1 (-.1)	-0.887
Hartford	23 (1.2)	37 (1.6)	41 (1.9)	-4 (-.3)	-1.842
McComb	23 (1.2)	38 (1.7)	40 (1.8)	-2 (-.1)	-0.897
Seattle	32 (1.5)	45 (2.1)	51 (2.4)	-6 (-.3)	-3.384
Portland	39 (1.7)	49 (2.3)	52 (2.5)	-3 (-.2)	-4.009
Jacksonville	23 (1.2)	44 (2.0)	36 (1.6)	+8 (+.4)	5.316
Hammond	31 (1.4)	46 (2.2)	45 (2.1)	+1 (+.1)	0.185
Bronx	22 (1.2)	40 (1.8)	43 (2.0)	-3 (-.2)	-2.016

TABLE R-1. (Continued)

GRADE 3-READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	45 (2.4)	67 (3.2)	58 (2.8)	+9 (+.4)	4.949
Athens	39 (2.3)	56 (2.7)	52 (2.6)	+4 (+.1)	1.746
Wichita	37 (2.2)	48 (2.5)	48 (2.5)	+0 (0)	0.043
Dallas	26 (1.8)	36 (2.2)	29 (1.9)	+7 (+.3)	3.661
Anchorage	42 (2.3)	62 (3.0)	63 (3.0)	-1 (0)	-0.239
Rockland	43 (2.4)	67 (3.2)	64 (3.1)	+3 (+.1)	1.517
Las Vegas	28 (1.9)	39 (2.3)	47 (2.4)	-8 (-.1)	-2.557
Fresno	42 (2.3)	54 (2.7)	55 (2.7)	-1 (0)	-0.654
Philadelphia	24 (1.7)	36 (2.2)	33 (2.1)	+3 (+.1)	1.252
Taft	40 (2.3)	54 (2.7)	53 (2.6)	-1 (+.1)	0.440
Grand Rapids	35 (2.2)	47 (2.4)	48 (2.5)	-1 (-.1)	-0.670
Hartford	31 (2.0)	51 (2.6)	50 (2.5)	+1 (+.1)	0.366
McComb	36 (2.2)	53 (2.6)	57 (2.8)	-4 (-.2)	-1.392
Seattle	46 (2.4)	53 (2.6)	64 (3.1)	-11 (-.5)	-3.457
Portland	41 (2.3)	58 (2.8)	61 (2.9)	-3 (-.1)	-1.127
Jacksonville	30 (2.0)	44 (2.4)	38 (2.2)	+6 (+.2)	3.106
Hammond	42 (2.3)	54 (2.7)	59 (2.8)	-5 (-.1)	-2.839
Bronx	31 (2.0)	47 (2.4)	53 (2.6)	-6 (-.2)	-2.199

TABLE R-1. (Continued)

GRADE 3-MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	59 (2.4)	88 (3.7)	75 (3.2)	+13 (+.5)	5.489
Athens	50 (2.3)	68 (2.9)	63 (2.6)	+5 (+.3)	2.172
Wichita	39 (1.9)	53 (2.3)	60 (2.5)	-7 (-.2)	-2.805
Dallas	35 (1.6)	59 (2.5)	39 (1.9)	+20 (+.6)	6.654
Anchorage	51 (2.3)	72 (3.0)	74 (3.1)	-2 (-.1)	-0.842
Rockland	57 (2.6)	79 (3.3)	80 (3.3)	-1 (0)	-0.606
Las Vegas	37 (1.7)	49 (2.3)	62 (2.6)	-13 (-.3)	-4.463
Fresno	51 (2.3)	62 (2.6)	68 (2.9)	-6 (-.3)	-2.896
Philadelphia	36 (1.7)	52 (2.3)	53 (2.3)	-1 (0)	-0.470
Taft	52 (2.3)	71 (3.0)	65 (2.7)	+6 (+.3)	4.130
Grand Rapids	42 (2.0)	59 (2.5)	61 (2.6)	-2 (-.1)	-1.042
Hartford	44 (2.2)	51 (2.3)	64 (2.7)	-13 (-.4)	-4.880
McComb	44 (2.2)	62 (2.6)	70 (3.0)	-8 (-.4)	-2.797
Seattle	53 (2.3)	71 (3.0)	78 (3.2)	-7 (-.2)	-2.391
Portland	58 (2.4)	74 (3.1)	80 (3.3)	-6 (-.2)	-2.792
Jacksonville	38 (1.8)	58 (2.4)	51 (2.3)	+7 (+.1)	3.906
Hammond	53 (2.3)	67 (2.8)	71 (3.0)	-4 (-.2)	-2.164
Bronx	42 (2.0)	65 (2.7)	69 (3.0)	-4 (-.3)	-1.555

TABLE R-1. (Continued)

GRADE 7-READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	54 (5.6)	61 (6.2)	59 (6.0)	+2 (+.2)	1.383
Athens	38 (4.4)	43 (4.8)	39 (4.5)	+4 (+.3)	2.089
Wichita	42 (4.8)	46 (4.8)	45 (4.8)	-1 (0)	0.733
Dallas	28 (3.5)	29 (3.9)	27 (3.7)	+2 (+.2)	0.844
Anchorage	53 (5.6)	59 (6.0)	59 (6.0)	+0 (0)	0.102
Rockland	56 (5.8)	59 (6.0)	64 (6.6)	-5 (-.6)	-2.418
Las Vegas	46 (5.0)	54 (5.6)	51 (5.5)	+3 (+.1)	1.637
Fresno	35 (4.2)	40 (4.5)	40 (4.5)	+0 (0)	-0.005
Philadelphia	24 (3.1)	28 (3.8)	29 (3.9)	-1 (-.1)	-0.712
Taft	36 (4.3)	40 (4.5)	41 (4.6)	-1 (-.1)	-0.911
Grand Rapids	37 (4.3)	41 (4.6)	43 (4.7)	-2 (-.1)	-0.890
Hartford	40 (4.5)	47 (4.9)	46 (4.8)	+1 (+.1)	0.270
McComb	35 (4.2)	42 (4.6)	40 (4.5)	+2 (+.1)	1.555
Seattle	57 (5.8)	62 (6.4)	61 (6.2)	+1 (+.2)	0.518
Portland	50 (5.3)	56 (5.8)	55 (5.7)	+1 (+.1)	0.838
Jacksonville	33 (3.9)	36 (4.3)	38 (4.4)	-2 (-.1)	-1.314
Hammond	50 (5.3)	57 (5.8)	56 (5.8)	+1 (0)	-0.626
Bronx	31 (3.8)	40 (4.5)	36 (4.3)	+4 (+.2)	0.987

TABLE R-1. (Continued)

GRADE 7-MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	63 (5.9)	79 (6.9)	76 (6.7)	+3 (+.2)	1.333
Athens	40 (4.5)	55 (5.4)	50 (5.1)	+5 (+.3)	1.651
Wichita	40 (4.5)	47 (4.9)	47 (4.9)	+0 (0)	-0.013
Dallas	30 (3.9)	34 (4.2)	35 (4.2)	-1 (0)	-0.229
Anchorage	53 (5.3)	63 (5.9)	60 (5.7)	+3 (+.2)	1.121
Rockland	62 (5.7)	72 (6.4)	79 (6.9)	-7 (-.5)	-3.203
Las Vegas	46 (4.8)	57 (5.4)	61 (5.7)	-4 (-.3)	-1.517
Fresno	36 (4.3)	42 (4.6)	52 (5.2)	-10 (-.6)	-4.200
Philadelphia	28 (3.8)	38 (4.4)	35 (4.2)	+3 (+.2)	1.302
Taft	46 (4.8)	52 (5.2)	53 (5.3)	-1 (-.1)	-0.528
Grand Rapids	38 (4.4)	44 (4.7)	50 (5.1)	-6 (-.4)	-3.499
Hartford	42 (4.6)	54 (5.3)	54 (5.3)	+0 (0)	-0.096
McComb	39 (4.5)	52 (5.2)	52 (5.2)	+0 (0)	0.030
Seattle	57 (5.4)	59 (5.6)	75 (6.7)	-16 (-1.1)	-5.532
Portland	53 (5.3)	65 (5.9)	65 (5.9)	+0 (0)	0.032
Jacksonville	38 (4.4)	46 (4.8)	47 (4.9)	-1 (-.1)	-0.518
Hammond	54 (5.3)	66 (6.0)	65 (5.9)	+1 (+.1)	0.810
Bronx	39 (4.5)	46 (4.8)	40 (4.5)	+6 (+.3)	1.218

TABLE R-1. (Continued)

GRADE 8-READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	49 (6.9)	55 (7.4)	55 (7.4)	+0 (0)	-0.148
Athens	27 (4.2)	30 (4.5)	33 (4.9)	-3 (-.4)	-1.857
Wichita	39 (5.7)	45 (6.4)	46 (6.6)	-1 (-.2)	-0.887
Dallas	23 (3.7)	31 (4.6)	27 (4.2)	+4 (+.4)	2.313
Anchorage	46 (6.6)	59 (8.0)	53 (7.3)	+6 (+.7)	2.647
Rockland	47 (6.7)	55 (7.4)	59 (8.0)	-4 (-.6)	-2.381
Las Vegas	39 (5.7)	45 (6.4)	43 (6.2)	+2 (+.2)	1.057
Fresno	30 (4.5)	33 (4.9)	38 (5.6)	-5 (-.7)	-2.507
Philadelphia	26 (4.1)	22 (3.5)	31 (4.6)	-9 (-1.1)	-4.938
Taft	30 (4.5)	36 (5.3)	36 (5.3)	+0 (0)	-0.250
Grand Rapids	32 (4.8)	42 (6.2)	35 (5.2)	+7 (+1.0)	3.984
Hartford	31 (4.6)	39 (5.7)	40 (5.8)	-1 (-.1)	-0.256
McComb	27 (4.2)	36 (5.3)	38 (5.6)	-2 (-.3)	-1.006
Seattle	41 (6.0)	43 (6.2)	53 (7.3)	-10 (-1.1)	-6.267
Portland	38 (5.6)	46 (6.6)	48 (6.7)	-2 (-.1)	-1.593
Jacksonville	29 (4.4)	34 (5.0)	33 (4.9)	+1 (+.1)	0.507
Hammond	40 (5.8)	47 (6.7)	51 (7.1)	-4 (-.4)	-2.639
Bronx	23 (3.7)	31 (4.6)	33 (4.9)	-2 (-.3)	-0.614

TABLE R-1. (Continued)

GRADE 8-MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	60 (7.1)	69 (8.1)	68 (8.0)	+1 (+.1)	0.580
Athens	32 (4.6)	35 (5.2)	37 (5.4)	-2 (-.2)	-1.064
Wichita	39 (5.4)	46 (6.2)	48 (6.4)	-2 (-.2)	-0.848
Dallas	26 (4.1)	37 (5.4)	34 (5.1)	+3 (-.3)	1.477
Anchorage	45 (5.9)	51 (6.7)	55 (6.9)	-4 (-.2)	-1.318
Rockland	56 (6.8)	66 (7.8)	70 (8.1)	-4 (-.3)	-2.027
Las Vegas	41 (5.6)	46 (6.2)	47 (6.4)	-1 (-.2)	-0.553
Fresno	34 (4.8)	37 (5.4)	42 (5.9)	-5 (-.5)	-2.760
Philadelphia	32 (4.6)	35 (5.2)	37 (5.4)	-2 (-.2)	-0.875
Taft	39 (5.4)	45 (6.2)	45 (6.2)	+0 (0)	-0.318
Grand Rapids	38 (5.3)	39 (5.7)	47 (6.4)	-8 (-.7)	-4.053
Hartford	39 (5.4)	43 (6.0)	45 (6.2)	-2 (-.2)	-1.484
McComb	39 (5.4)	47 (6.4)	49 (6.6)	-2 (-.2)	-1.240
Seattle	47 (6.0)	44 (6.0)	64 (7.6)	-20 (-1.6)	-9.367
Portland	47 (6.0)	52 (6.7)	60 (7.3)	-8 (-.6)	-3.657
Jacksonville	38 (5.3)	43 (6.0)	44 (6.0)	-1 (0)	-0.317
Hammond	46 (6.0)	51 (6.7)	55 (6.9)	-4 (-.2)	-2.454
Bronx	30 (4.4)	35 (5.2)	38 (5.6)	-3 (-.4)	-0.793

TABLE R-1. (Continued)

GRADE 9-READING

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	48 (6.7)	55 (7.4)	53 (7.3)	+2 (+.1)	0.864
Athens	34 (5.0)	43 (6.2)	39 (5.7)	+4 (+.5)	2.290
Wichita	48 (6.7)	52 (7.3)	54 (7.4)	-2 (-.1)	-0.774
Dallas	31 (4.6)	36 (5.3)	36 (5.3)	+0 (0)	-0.244
Anchorage	55 (7.4)	61 (8.2)	56 (7.6)	+5 (+.6)	1.749
Rockland	59 (8.0)	66 (9.3)	65 (9.0)	+1 (+.3)	0.836
Las Vegas	49 (6.9)	58 (8.0)	52 (7.3)	+6 (+.7)	2.412
Fresno	31 (4.6)	39 (5.7)	36 (5.3)	+3 (+.4)	1.191
Philadelphia	23 (3.7)	24 (3.8)	27 (4.2)	-3 (-.4)	-2.23
Taft	36 (5.3)	45 (6.4)	37 (5.5)	+8 (+.9)	3.723
Grand Rapids	37 (5.5)	44 (6.4)	41 (6.0)	+3 (+.4)	1.496
Hartford	32 (4.8)	39 (5.7)	39 (5.7)	+0 (0)	0.0861
McComb	35 (5.2)	41 (6.0)	42 (6.2)	-1 (-.2)	-0.880
Seattle	49 (6.9)	54 (7.4)	58 (8.0)	-4 (-.6)	-2.030
Portland	47 (6.9)	53 (7.3)	56 (7.6)	-3 (-.3)	-2.006
Jacksonville	33 (4.9)	41 (6.0)	39 (5.7)	+2 (+.3)	1.256
Hammond	52 (7.3)	58 (8.0)	59 (8.0)	-1 (0)	-0.424
Bronx	No Data	No Data	No Data	No Data	No Data

TABLE R-1. (Continued)

GRADE 9-MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/C Difference	t Ratio
		E	C		
Selmer	59 (6.9)	68 (8.0)	73 (8.4)	-5 (-.4)	-2.412
Athens	41 (5.6)	53 (6.8)	48 (6.4)	+5 (+.4)	2.341
Wichita	48 (6.2)	55 (6.9)	58 (7.1)	-3 (-.2)	-1.249
Dallas	38 (5.3)	45 (6.2)	44 (6.0)	+1 (+.2)	0.180
Anchorage	53 (6.6)	65 (7.8)	55 (6.9)	+10 (+.9)	4.061
Rockland	70 (7.8)	78 (9.0)	70 (8.1)	+8 (+.9)	3.473
Las Vegas	53 (6.6)	62 (7.5)	62 (7.5)	+0 (0)	0.123
Fresno	37 (5.2)	40 (5.7)	40 (5.7)	+0 (0)	0.440
Philadelphia	31 (4.5)	33 (4.9)	37 (5.4)	-4 (-.5)	-1.954
Taft	48 (6.2)	55 (6.9)	51 (6.7)	+4 (+.2)	1.509
Grand Rapids	46 (6.0)	47 (6.4)	51 (6.7)	-4 (-.3)	-1.672
Hartford	41 (5.6)	47 (6.4)	44 (6.0)	+3 (+.4)	1.572
McComb	48 (6.2)	55 (6.9)	56 (6.9)	-1 (0)	-0.667
Seattle	54 (6.7)	60 (7.3)	71 (8.0)	-11 (-.7)	-4.668
Portland	54 (6.7)	61 (7.5)	67 (8.0)	-6 (-.5)	-2.806
Jacksonville	42 (5.7)	49 (6.6)	48 (6.4)	+1 (+.2)	0.543
Hammond	58 (6.9)	70 (8.1)	71 (8.2)	-1 (-.1)	-0.891
Bronx	No Data	No Data	No Data	No Data	No Data

APPENDIX S

S-1

TABLE S-1 . REGRESSION RESULTS FOR SECOND-YEAR TESTING

A. GRADE 1 - READING

Site	Mean Pretest Value*	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Selmer	(3)	81 (2.2)	73 (1.8)	+8 (+.4)	2.574
Athens	(2)	61 (1.1)	65 (1.4)	-4 (-0.3)	-1.482
Dallas	(1)	63 (1.2)	73 (1.8)	-10 (-0.6)	-2.219
Anchorage	(4)	70 (1.7)	84 (2.4)	-14 (-.7)	-1.852
Jacksonville	(2)	66 (1.5)	57 (0.8)	9 (0.7)	3.859

B. GRADE 1 - MATHEMATICS

Site	Mean Pretest Value*	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Athens	(2)	47 (1.7)	53 (1.9)	-6 (-0.2)	-2.080
Dallas	(1)	49 (1.6)	55 (2.0)	-6 (-0.4)	-1.156
Anchorage	(4)	51 (1.9)	64 (2.1)	-13 (-0.2)	-1.878
Fresno	(4)	48 (1.5)	53 (1.7)	-5 (-.2)	-1.575
Portland	(4)	58 (1.9)	62 (2.0)	-4 (-.1)	-1.785
Jacksonville	(2)	46 (1.7)	34 (1.0)	12 (0.3)	5.415

* The stanine value corresponding to the combined E and C pretest raw score mean.

TABLE S-1. (Continued)

C. GRADE 2 - READING

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Selmer	40 (1.7)	69 (2.4)	68 (2.4)	+1 (.0)	0.776
Anchorage	36 (1.6)	66 (2.3)	68 (2.4)	-2 (-.1)	-0.469
Las Vegas	23 (1.3)	44 (1.8)	53 (2.0)	-9 (-.2)	-1.986
Jacksonville	28 (1.4)	57 (2.1)	51 (2.0)	+6 (+.1)	2.190

D. GRADE 2 - MATHEMATICS

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Selmer	36 (1.6)	55 (3.0)	54 (2.7)	+1 (+.3)	0.950
Dallas	23 (1.2)	44 (2.0)	31 (1.4)	+13 (+.6)	3.985
Anchorage	32 (1.5)	55 (3.0)	54 (2.7)	+1 (+.3)	0.752
Las Vegas	21 (1.2)	37 (1.6)	45 (2.1)	-8 (-.5)	-2.799
Fresno	32 (1.5)	49 (2.3)	53 (2.6)	-4 (-.3)	-3.619
Portland	40 (1.8)	54 (2.7)	56 (3.1)	-2 (-.4)	-2.319
Jacksonville	22 (1.2)	46 (2.2)	40 (1.8)	+6 (+.4)	3.187

TABLE S-1. (Continued)

E. GRADE 3 - READING

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Selmer	45 (2.4)	72 (3.4)	66 (3.1)	+6 (+.3)	3.188
Athens	39 (2.3)	61 (2.9)	57 (2.8)	+4 (+.1)	1.746
Dallas	27 (1.8)	40 (2.3)	34 (2.2)	+6 (+.1)	2.052
Anchorage	43 (2.4)	66 (3.1)	67 (3.2)	-1 (-.1)	-0.265
Las Vegas	29 (1.9)	47 (2.4)	56 (2.7)	-9 (-.3)	-2.481
Jacksonville	30 (2.0)	48 (2.5)	44 (2.4)	+4 (+.1)	1.672

F. GRADE 3 - MATHEMATICS

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Selmer	58 (2.4)	90 (3.9)	79 (3.3)	+11 (+.6)	4.983
Athens	51 (2.3)	78 (3.2)	72 (3.0)	+6 (+.2)	2.502
Dallas	36 (1.7)	57 (2.4)	49 (2.3)	+8 (+.1)	2.413
Las Vegas	37 (1.7)	61 (2.6)	68 (2.9)	-7 (-.3)	-1.359
Fresno	51 (2.3)	76 (3.2)	76 (3.2)	0 (0)	-0.000
Portland	58 (2.4)	81 (3.4)	85 (3.5)	-4 (-.1)	-1.935
Jacksonville	37 (1.7)	63 (2.6)	55 (2.4)	+8 (+.2)	3.835

TABLE S-1. (Continued)

G. GRADE 7 - READING

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Athens	39 (4.5)	45 (4.9)	41 (4.6)	+4 (+.3)	1.419
Dallas	28 (3.5)	33 (3.9)	31 (3.8)	+2 (+.1)	0.738

H. GRADE 7 - MATHEMATICS

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Athens	41 (4.6)	54 (5.3)	47 (4.9)	+7 (+.4)	2.286
Grand Rapids	39 (4.5)	52 (5.2)	56 (5.4)	-4 (-.2)	-1.547

I. GRADE 8 - READING

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Dallas	23 (3.7)	31 (4.6)	27 (4.2)	+4 (+.4)	1.519
Anchorage	46 (6.6)	44 (6.4)	49 (6.9)	-5 (-.5)	-1.623
Grand Rapids	32 (4.8)	43 (6.2)	37 (5.5)	+6 (+.7)	2.493

J. GRADE 8 - MATHEMATICS

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Dallas	26 (4.1)	40 (5.7)	39 (5.7)	+1 (0)	0.414
Grand Rapids	37 (5.2)	50 (6.6)	53 (6.8)	-.3 (-.2)	-.888

TABLE S-1. (Continued)

K. GRADE 9 - READING

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Grand Rapids	38 (5.6)	50 (7.1)	45 (6.4)	+5 (+.7)	1.737

L. GRADE 9 - MATHEMATICS

Site	Mean Pretest Value	Estimated Second-Year Test Means		E/C Difference	t Ratio
		E	C		
Athens	39 (5.4)	55 (6.9)	51 (6.7)	+4 (+.2)	1.340

APPENDIX T

REGRESSION RESULTS FOR CONTROL VERSUS COMPARISON GROUPS

TABLE T-1. REGRESSION RESULTS FOR CONTROL (C) VERSUS COMPARISON (R) GROUPS

A. GRADE 1 - READING

Site	Mean Pretest Value*	Estimated Posttest Means C	Estimated Posttest Means R	C/R Difference	t Ratio
Wichita	(3)	73 (1.8)	70 (1.7)	3 (0.1)	0.871
Rockland	(5)	80 (2.2)	86 (2.5)	-6 (-0.3)	-2.499
McComb	(3)	70 (1.7)	65 (1.4)	5 (0.3)	1.611
Jacksonville	(2)	35 (<.6)	56 (0.8)	-21 (<-0.2)	-8.080
Hammond	(3)	75 (1.9)	62 (1.2)	13 (0.7)	4.942

B. GRADE 1 - MATHEMATICS

Site	Mean Pretest Value*	Estimated Posttest Means C	Estimated Posttest Means R	C/R Difference	t Ratio
Wichita	(3)	42 (1.5)	47 (1.7)	-5 (-0.2)	-1.822
Rockland	(5)	59 (2.2)	62 (2.4)	-3 (-0.2)	-1.287
McComb	(3)	43 (1.5)	40 (1.3)	3 (0.2)	0.832
Jacksonville	(2)	23 (<.6)	33 (0.9)	-10 (<-0.3)	-5.070
Hammond	(3)	48 (1.7)	49 (1.8)	-1 (-0.1)	-0.377

* The stanine value corresponding to the combined C and R pretest raw score mean.

TABLE T-1. (Continued)

C. GRADE 2 - READING

Site	Mean Pretest Value	Estimated Posttest Means C	Estimated Posttest Means R	C/R Difference	t Ratio
Wichita	44 (1.8)	61 (2.1)	64 (2.2)	-3 (-.1)	-1.264
Rockland	61 (2.2)	74 (3.1)	74 (3.1)	0 (0)	-0.666
McComb	40 (1.7)	63 (2.2)	59 (2.1)	+4 (+.1)	1.313
Jacksonville	28 (1.4)	44 (1.8)	47 (1.8)	-3 (0)	-1.067
Hammond	38 (1.7)	58 (2.0)	58 (2.0)	0 (0)	0.368

D. GRADE 2 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means C	Estimated Posttest Means R	C/R Difference	t Ratio
Wichita	33 (1.5)	44 (2.0)	46 (2.2)	-2 (-.2)	-0.821
Rockland	45 (2.1)	58 (3.5)	59 (3.8)	-1 (-.3)	-1.887
McComb	31 (1.4)	47 (2.2)	47 (2.2)	0 (0)	0.193
Jacksonville	25 (1.3)	36 (1.6)	44 (2.0)	-8 (-.4)	-4.161
Hammond	34 (1.5)	48 (2.3)	45 (2.1)	+3 (+.2)	1.636

TABLE T-1. (Continued)

E. GRADE 3 - READING

Site	Mean Pretest Value	Estimated Posttest Means		C/R Difference	t Ratio
		C	R		
Athens	50 (2.4)	63 (3.0)	64 (3.1)	-1 (-.1)	-0.232
Wichita	46 (2.4)	54 (2.7)	60 (2.9)	-6 (-.2)	-1.982
Rockland	64 (2.8)	75 (3.6)	76 (3.7)	-1 (-.1)	-1.384
McComb	47 (2.4)	65 (3.1)	59 (2.8)	+6 (+.3)	2.618
Jacksonville	33 (2.1)	42 (2.4)	47 (2.4)	-5 (0)	-2.335
Hammond	48 (2.4)	64 (3.1)	59 (2.8)	+5 (+.2)	2.553

F. GRADE 3 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		C/R Difference	t Ratio
		C	R		
Athens	58 (2.4)	72 (3.0)	86 (3.5)	-14 (-.5)	-4.870
Wichita	44 (2.2)	63 (2.6)	63 (2.6)	0 (0)	-0.069
Rockland	71 (2.9)	89 (3.8)	92 (4.1)	-3 (-.3)	-2.342
McComb	55 (2.4)	81 (3.4)	72 (3.0)	+9 (+.4)	3.304
Jacksonville	38 (1.8)	52 (2.3)	58 (2.4)	-6 (-.1)	-2.690
Hammond	57 (2.4)	75 (3.2)	74 (3.1)	+1 (+.1)	0.405

TABLE T-1. (Continued)

G. GRADE 7 - READING

Site	Mean Pretest Value	Estimated Posttest Means		C/R Difference	t Ratio
		C	R		
Athens	49 (5.3)	48 (5.2)	59 (6.0)	-11 (-.8)	-4.295
Wichita	44 (4.9)	47 (5.0)	47 (5.0)	0 (0)	0.058
Rockland	73 (7.4)	78 (8.2)	79 (8.4)	-1 (-.2)	-0.815
Hartford	37 (4.3)	43 (4.8)	42 (4.8)	+1 (0)	0.402
Jacksonville	34 (4.1)	39 (4.5)	38 (4.4)	+1 (+.1)	0.283

H. GRADE 7 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		C/R Difference	t Ratio
		C	R		
Athens	51 (5.2)	61 (5.7)	68 (6.2)	-7 (-.5)	-2.238
Wichita	42 (4.6)	50 (5.1)	49 (5.1)	+1 (0)	0.122
Rockland	79 (6.9)	92 (8.2)	94 (8.6)	-2 (-.4)	-1.496
Hartford	40 (4.5)	52 (5.2)	56 (5.4)	-4 (-.2)	-1.378
Jacksonville	39 (4.5)	48 (4.9)	41 (4.6)	+7 (+.3)	2.178

TABLE T-1 (Continued)

I. GRADE 8 - READING

Site	Mean Pretest Value	Estimated Posttest Means		C/R Difference	t Ratio
		C	R		
Athens	34 (5.0)	38 (5.6)	44 (6.4)	-6 (-.8)	-2.259
Wichita	43 (6.2)	50 (7.1)	50 (7.1)	0 (0)	0.009
Rockland	66 (9.3)	72 (10.4)	74 (11.0)	-2 (-.6)	-0.945
Hartford	28 (4.3)	37 (5.5)	36 (5.3)	+1 (+.2)	0.853

J. GRADE 8 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		C/R Difference	t Ratio
		C	R		
Athens	37 (5.2)	41 (5.9)	60 (7.3)	-19 (-1.4)	-5.596
Wichita	44 (5.9)	51 (6.7)	55 (6.9)	-4 (-.2)	-1.062
Rockland	76 (8.2)	86 (10.1)	87 (10.4)	-1 (-.3)	-0.568
Hartford	35 (4.9)	41 (5.9)	42 (5.9)	-1 (-.3)	-0.627

TABLE T-1 (Continued)

K. GRADE 9 - READING

Site	Mean Pretest Value	Estimated Posttest Means C	Estimated Posttest Means R	C/R Difference	t Ratio
Athens	44 (6.4)	48 (6.7)	53 (7.3)	-5 (-.6)	-1.812
Wichita	52 (7.3)	57 (7.8)	56 (7.6)	+1 (+.2)	0.223
Rockland	73 (10.7)	76 (11.6)	76 (11.6)	0 (0)	-0.157
Jacksonville	33 (4.9)	39 (5.7)	40 (5.8)	-1 (-.1)	-0.427

L. GRADE 9 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means C	Estimated Posttest Means R	C/R Difference	t Ratio
Athens	53 (6.6)	60 (7.3)	75 (8.6)	-15 (-1.3)	-4.575
Wichita	53 (6.6)	62 (7.5)	56 (6.9)	+6 (+.6)	1.683
Rockland	88 (9.8)	90 (11.0)	93 (11.6)	-3 (-.6)	-2.633
Jacksonville	44 (5.9)	50 (6.6)	47 (6.4)	+3 (+.2)	0.751

APPENDIX U

REGRESSION RESULTS FOR EXPERIMENTAL VERSUS COMPARISON GROUPS

113

TABLE U-1. REGRESSION RESULTS FOR EXPERIMENTAL (E) VERSUS COMPARISON (R) GROUPS

A. GRADE 1 - READING

Site	Mean Pretest Value*	Estimated Posttest Means E	Estimated Posttest Means R	E/R Difference	t Ratio
Wichita	(3)	55 (0.7)	69 (1.6)	-14 (-0.9)	-3.316
Rockland	(5)	74 (1.9)	83 (2.3)	-9 (-0.4)	-3.548
McComb	(2)	59 (1.0)	60 (1.0)	-1 (0)	-0.152
Jacksonville	(2)	58 (0.9)	59 (1.0)	-1 (-0.1)	-0.577
Hammond	(3)	62 (1.2)	58 (0.9)	4 (0.3)	1.063

B. GRADE 1 - MATHEMATICS

Site	Mean Pretest Value*	Estimated Posttest Means E	Estimated Posttest Means R	E/R Difference	t Ratio
Wichita	(3)	38 (1.2)	47 (1.7)	-9 (-0.5)	-2.842
Rockland	(5)	54 (2.0)	59 (2.2)	-5 (-0.2)	-2.140
McComb	(2)	44 (1.6)	35 (1.1)	-9 (0.5)	2.016
Jacksonville	(2)	39 (1.3)	35 (1.1)	4 (0.2)	1.923
Hammond	(3)	45 (1.6)	46 (1.7)	-1 (-0.1)	-0.529

* The stanine value corresponding to the combined E and R pretest raw score mean.

TABLE U-1 (Continued)

C. GRADE 2 - READING

Site	Mean Pretest Value	Estimated Posttest Means E	Estimated Posttest Means R	E/R Difference	t Ratio
Wichita	37 (1.6)	61 (2.1)	60 (2.0)	+1 (+.1)	0.355
Rockland	51 (2.0)	71 (2.5)	74 (3.1)	-3 (-.6)	-0.998
McComb	32 (1.5)	50 (1.9)	52 (1.9)	-2 (0)	-0.576
Jacksonville	30 (1.4)	53 (2.0)	50 (1.9)	+3 (+.1)	1.072
Hammond	35 (1.6)	54 (2.0)	55 (2.0)	-1 (0)	-0.538

D. GRADE 2 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means E	Estimated Posttest Means R	E/R Difference	t Ratio
Wichita	30 (1.4)	42 (1.9)	44 (2.0)	-2 (-.1)	-1.275
Rockland	39 (1.7)	53 (2.6)	58 (3.5)	-5 (-.9)	-3.458
McComb	26 (1.3)	40 (1.8)	44 (2.0)	-4 (-.2)	-1.624
Jacksonville	24 (1.3)	44 (2.0)	43 (2.0)	+1 (0)	0.967
Hammond	33 (1.5)	46 (2.2)	45 (2.1)	+1 (+.1)	0.608

1/5

TABLE U-1. (Continued)

E. GRADE 3 - READING

Site	Mean Pretest Value	Estimated Posttest Means		E/R Difference	t Ratio
		E	R		
Athens	44 (2.4)	60 (2.9)	60 (2.9)	+0 (0)	0.089
Wichita	38 (2.2)	49 (2.5)	53 (2.6)	-4 (-.1)	-1.884
Rockland	54 (2.5)	78 (4.1)	74 (3.5)	+4 (+.6)	1.419
McComb	32 (2.1)	48 (2.5)	47 (2.4)	+1 (+.1)	0.231
Jacksonville	33 (2.1)	47 (2.4)	47 (2.4)	0 (0)	0.054
Hammond	42 (2.3)	55 (2.7)	54 (2.7)	+1 (0)	0.124

F. GRADE 3 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/R Difference	t Ratio
		E	R		
Athens	51 (2.3)	69 (3.0)	83 (3.5)	-14 (-.5)	-3.725
Wichita	46 (2.2)	58 (2.4)	64 (2.7)	-6 (-.3)	-2.643
Rockland	63 (2.6)	85 (3.5)	88 (3.7)	-3 (-.2)	-1.610
McComb	41 (2.0)	59 (2.5)	59 (2.5)	0 (0)	0.010
Jacksonville	39 (1.9)	60 (2.5)	58 (2.4)	+2 (+.1)	0.723
Hammond	51 (2.3)	65 (2.7)	68 (2.9)	-3 (-.2)	-1.220

TABLE U-1. (Continued)

G. GRADE 7 - READING

Site	Mean Pretest Value	Estimated Posttest Means		E/R Difference	t Ratio
		E	R		
Athens	46 (5.0)	51 (5.5)	56 (5.8)	-5 (-.3)	-1.571
Wichita	45 (4.9)	50 (5.3)	48 (5.2)	+2 (+.1)	0.608
Rockland	65 (6.6)	65 (6.6)	74 (7.6)	-9 (-1.0)	-4.189
Hartford	35 (4.2)	42 (4.8)	41 (4.6)	+1 (+.2)	0.780
Jacksonville	31 (3.8)	34 (4.1)	36 (4.3)	-2 (-.2)	-0.723

H. GRADE 7 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/R Difference	t Ratio
		E	R		
Athens	44 (4.7)	59 (5.6)	62 (5.7)	-3 (-.1)	-0.497
Wichita	44 (4.7)	52 (5.2)	52 (5.2)	+0 (0)	0.050
Rockland	70 (6.2)	80 (6.9)	88 (7.8)	-8 (-.9)	-4.810
Hartford	38 (4.4)	50 (5.1)	54 (5.3)	-4 (-.2)	-1.598
Jacksonville	35 (4.2)	44 (4.7)	40 (4.5)	+4 (+.2)	1.343

TABLE U-1. (Continued)

I. GRADE 8 - READING

Site	Mean Pretest Value	Estimated Posttest Means		E/R Difference	t Ratio
		E	R		
Athens	32 (4.8)	32 (4.8)	43 (6.2)	-11 (-1.4)	-3.042
Wichita	39 (5.7)	44 (6.4)	48 (7.6)	-4 (-.3)	-1.187
Rockland	58 (8.0)	64 (8.7)	68 (9.6)	-4 (-.9)	-2.129
Hartford	31 (4.6)	39 (5.7)	38 (5.6)	+1 (+.1)	0.815

J. GRADE 8 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means		E/R Difference	t Ratio
		E	R		
Athens	38 (5.3)	38 (5.6)	61 (7.5)	-23 (-1.9)	-5.951
Wichita	38 (5.3)	45 (6.2)	51 (6.7)	-6 (-.5)	-1.372
Rockland	67 (7.6)	77 (8.8)	80 (9.2)	-3 (-.4)	-1.330
Hartford	38 (5.3)	42 (5.9)	44 (6.0)	-2 (-.1)	-1.206

TABLE U-1. (Continued)

K. GRADE 9 - READING

Site	Mean Pretest Value	Estimated Posttest Means E	Estimated Posttest Means R	E/R Difference	t Ratio
Athens	39 (5.7)	50 (7.1)	49 (6.9)	+1 (+.2)	0.053
Wichita	47 (6.7)	51 (7.1)	53 (7.3)	-2 (-.2)	-0.542
Rockland	64 (8.7)	70 (9.8)	68 (9.6)	+2 (+.2)	1.638
Jacksonville	33 (4.9)	41 (6.0)	40 (5.8)	+1 (+.2)	0.388

L. GRADE 9 - MATHEMATICS

Site	Mean Pretest Value	Estimated Posttest Means E	Estimated Posttest Means R	E/R Difference	t Ratio
Athens	49 (6.2)	62 (7.5)	72 (8.2)	-10 (-.7)	-2.009
Wichita	46 (6.0)	54 (6.8)	52 (6.7)	+2 (+.1)	0.571
Rockland	74 (8.1)	81 (9.2)	81 (9.2)	0 (0)	0.291
Jacksonville	43 (5.7)	49 (6.6)	46 (6.2)	+3 (+.4)	0.985

APPENDIX V

REGRESSION RESULTS FOR SPECIAL-TREATMENT GROUPS AT GRAND RAPIDS

TABLE V-1. REGRESSION RESULTS IN READING FOR EXPERIMENTAL (E) VERSUS SPECIAL TREATMENT (ST) GROUPS AT GRAND RAPIDS

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Westinghouse	(2)	52 (<.6)	55 (0.7)	-3 (<-0.1)	-0.816
Project Read	(2)	52 (<.6)	59 (1.0)	-7 (<-0.4)	-2.066

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Reading Centers	26 (1.3)	46 (1.8)	58 (2.0)	-12 (-.2)	-2.785
Westinghouse	25 (1.3)	45 (1.8)	46 (1.8)	-1 (0)	-0.233
Project Read	28 (1.4)	48 (1.9)	50 (1.9)	-2 (0)	-0.715

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Reading Centers	31 (2.0)	43 (2.4)	52 (2.6)	-9 (-.2)	-2.731
Westinghouse	34 (2.2)	45 (2.4)	51 (2.6)	-6 (-.2)	-2.084
Project Read	29 (1.9)	41 (2.3)	45 (2.4)	-4 (-.1)	-1.383

* The stanine value corresponding to the combined E and ST pretest raw score mean.

TABLE V-1. (Continued)

D. GRADE 7

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Reading Centers	34 (4.1)	39 (4.5)	36 (4.3)	+3 (+.2)	1.202
COMES	33 (3.9)	38 (4.4)	42 (4.8)	-4 (-.4)	-2.423

E. GRADE 8

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
COMES	26 (4.1)	37 (5.5)	34 (5.0)	+3 (+.5)	1.285

F. GRADE 9

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
COMES	32 (4.8)	39 (5.7)	34 (5.0)	+5 (+.7)	1.570

TABLE V-2. REGRESSION RESULTS IN MATHEMATICS FOR EXPERIMENTAL (E) VERSUS SPECIAL TREATMENT (ST) GROUPS AT GRAND RAPIDS

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Westinghouse	(2)	31 (0.8)	32 (0.9)	-1 (-0.1)	-0.464
Project Read	(2)	30 (0.7)	37 (1.2)	-7 (-0.5)	-2.239

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Reading Centers	32 (1.5)	44 (2.0)	49 (2.3)	-5 (-.3)	-2.795
Westinghouse	31 (1.4)	43 (2.0)	44 (2.0)	-1 (0)	-0.301
Project Read	32 (1.5)	44 (2.0)	46 (2.2)	-2 (-.2)	-1.835

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Reading Centers	40 (1.9)	57 (2.4)	65 (2.7)	-8 (-.3)	-2.781
Westinghouse	41 (2.0)	58 (2.4)	58 (2.4)	0 (0)	-0.110
Project Read	38 (1.8)	56 (2.4)	67 (2.8)	-11 (-.4)	-4.493

* The stanine value corresponding to the combined E and ST pretest raw score mean.

TABLE V-2. (Continued)

D. GRADE 7

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Reading Centers	37 (4.3)	43 (4.7)	45 (4.8)	-2 (-.1)	-0.650
COMES	33 (4.1)	41 (4.6)	45 (4.8)	-4 (-.2)	-2.312

E. GRADE 8

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
COMES	35 (4.9)	37 (5.4)	36 (5.3)	+1 (+.1)	0.433

F. GRADE 9

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
COMES	44 (5.9)	47 (6.4)	40 (5.7)	+7 (+.7)	1.209

TABLE V-3. REGRESSION RESULTS IN READING FOR CONTROL (C) VERSUS SPECIAL-TREATMENT (ST) GROUPS AT GRAND RAPIDS

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Westinghouse	(3)	60 (1.0)	60 (1.0)	0 (0)	-0.040
Project Read	(3)	60 (1.0)	69 (1.6)	-9 (-0.6)	-2.542

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Reading Centers	34 (1.5)	56 (2.0)	63 (2.2)	-7 (-.2)	-1.797
Westinghouse	33 (1.5)	56 (2.0)	53 (2.0)	+3 (0)	0.702
Project Read	35 (1.6)	57 (2.0)	55 (2.0)	+2 (0)	0.505

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Reading Centers	34 (2.2)	47 (2.4)	53 (2.6)	-6 (-.2)	-1.498
Westinghouse	37 (2.2)	50 (2.5)	53 (2.6)	-3 (-.1)	-1.442
Project Read	32 (2.1)	45 (2.4)	46 (2.4)	-1 (0)	-0.161

* The stanine value corresponding to the combined C and ST pretest raw score mean.

TABLE V-3. (Continued)

D. GRADE 7

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Reading Centers	40 (4.5)	45 (4.9)	41 (4.6)	+4 (+.3)	1.476
COMES	36 (4.3)	42 (4.8)	45 (4.9)	-3 (-.1)	-1.914

E. GRADE 8

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
COMES	31 (4.6)	34 (5.0)	39 (5.7)	-5 (-.7)	-2.018

F. GRADE 9

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
COMES	35 (5.2)	39 (5.7)	35 (5.2)	+4 (+.5)	1.094

TABLE V-4. REGRESSION RESULTS IN MATHEMATICS FOR CONTROL (C) VERSUS SPECIAL-TREATMENT (ST) GROUPS AT GRAND RAPIDS

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means C	Estimated Posttest Means ST	C/ST Difference	t Ratio
Westinghouse	(3)	43 (1.5)	36 (1.1)	7 (0.4)	1.867
Project Read	(3)	43 (1.5)	45 (1.6)	-2 (-0.1)	-0.531

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means C	Estimated Posttest Means ST	C/ST Difference	t Ratio
Reading Centers	31 (1.4)	44 (2.0)	48 (2.3)	-4 (-.3)	-1.885
Westinghouse	30 (1.4	44 (2.0)	43 (2.0)	+1 (0)	0.404
Project Read	31 (1.4)	44 (2.0)	46 (2.2)	-2 (-.2)	-0.726

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means C	Estimated Posttest Means ST	C/ST Difference	t Ratio
Reading Centers	43 (2.1)	62 (2.6)	68 (2.9)	-6 (-.3)	-2.004
Westinghouse	44 (2.2)	62 (2.6)	61 (2.6)	+1 (0)	0.470
Project Read	40 (1.9)	59 (2.5)	67 (2.8)	-8 (-.3)	-2.918

* The stanine value corresponding to the combined C and ST pretest raw score mean.

TABLE V-4. (Continued)

D. GRADE 7

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Reading Centers	41 (4.6)	53 (5.3)	48 (4.9)	+5 (-.4)	1.687
COMES	36 (4.3)	48 (4.9)	47 (4.9)	+1 (0)	0.397

E. GRADE 8

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
COMES	36 (5.1)	46 (6.2)	37 (5.4)	+9 (+.8)	3.126

F. GRADE 9

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
COMES	41 (5.6)	46 (6.2)	38 (5.6)	+8 (+.6)	1.903

APPENDIX W

REGRESSION RESULTS FOR SPECIAL-TREATMENT GROUPS AT HARTFORD

129

TABLE W-1. REGRESSION RESULTS IN READING FOR EXPERIMENTAL (E) VERSUS SPECIAL-TREATMENT (ST) GROUPS AT HARTFORD

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Waverly	(2)	54 (0.7)	65 (1.4)	-11 (-0.7)	-3.853
Project CONCERN	(2)	54 (0.7)	52 (<.6)	2 (>0.1)	0.404

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Waverly	26 (1.3)	42 (1.7)	52 (1.9)	-10 (-.2)	-3.803
Project CONCERN	29 (1.4)	43 (1.7)	53 (2.0)	-10 (-.3)	-2.928

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Waverly	29 (1.9)	49 (2.5)	52 (2.6)	-3 (-.1)	-1.059
Project CONCERN	29 (1.9)	49 (2.5)	51 (2.6)	-2 (-.1)	-0.336

* The stanine value corresponding to the combined E and ST pretest raw score mean.

TABLE W-2. REGRESSION RESULTS IN MATHEMATICS FOR EXPERIMENTAL (E) VERSUS SPECIAL-TREATMENT (ST) GROUPS AT HARTFORD

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Waverly	(2)	33 (0.9)	44 (1.6)	-11 (-0.7)	-4.819
Project CONCERN	(2)	33 (0.9)	33 (0.9)	0 (0)	-0.217

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Waverly	22 (1.2)	36 (1.6)	46 (2.2)	-10 (-.6)	-6.470
Project CONCERN	24 (1.3)	38 (1.7)	40 (1.8)	-2 (-.1)	-0.696

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		E/ST Difference	t Ratio
		E	ST		
Waverly	40 (1.9)	48 (2.3)	69 (3.0)	-21 (-.7)	-6.757
Project CONCERN	41 (2.0)	49 (2.3)	58 (2.4)	-9 (-.1)	-1.576

* The stanine value corresponding to the combined E and ST pretest raw score mean.

TABLE W-3. REGRESSION RESULTS IN READING FOR CONTROL (C) VERSUS SPECIAL-TREATMENT (ST) GROUPS AT HARTFORD

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Waverly	(2)	58 (0.9)	65 (1.4)	-7 (-0.5)	-3.086
Project CONCERN	(2)	58 (0.9)	51 (<.6)	7 (>0.3)	1.470

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Waverly	30 (1.4)	53 (2.0)	54 (2.0)	-1 (0)	-0.434
Project CONCERN	33 (1.5)	56 (2.0)	55 (2.0)	+1 (0)	0.151

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Waverly	35 (2.2)	53 (2.6)	55 (2.7)	-2 (-.1)	-0.718
Project CONCERN	41 (2.3)	56 (2.7)	56 (2.7)	0 (0)	0.032

* The stanine value corresponding to the combined C and ST pretest raw score mean.

TABLE W-4. REGRESSION RESULTS IN MATHEMATICS FOR CONTROL (C) VERSUS SPECIAL-TREATMENT (ST) GROUPS AT HARTFORD

A. GRADE 1

Program	Mean Pretest Value*	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Waverly	(2)	41 (1.4)	44 (1.6)	-3 (-0.2)	-1.226
Project CONCERN	(2)	41 (1.4)	33 (0.9)	8 (0.5)	2.739

B. GRADE 2

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Waverly	26 (1.3)	42 (1.9)	48 (2.3)	-6 (-.4)	-3.900
Project CONCERN	29 (1.4)	44 (2.0)	43 (2.0)	+1 (0)	0.475

C. GRADE 3

Program	Mean Pretest Value	Estimated Posttest Means		C/ST Difference	t Ratio
		C	ST		
Waverly	42 (2.0)	62 (2.6)	70 (3.0)	-8 (-.4)	-2.383
Project CONCERN	47 (2.2)	66 (2.8)	63 (2.6)	+3 (+.2)	0.375

* The stanine value corresponding to the combined C and ST pretest raw score mean.

APPENDIX X

REGRESSION RESULTS FOR ATTENDANCE

TABLE X-1. REGRESSION RESULTS FOR ATTENDANCE

A. GRADE 1

Grade	Median 69-70 Attendance*	Estimated Mean 70-71 Attendance*		E/C Difference*	t Ratio
		E	C		
Rockland	6.55	6.14	5.29	+0.85	1.144

B. GRADE 2

Grade	Median 69-70 Attendance*	Estimated Mean 70-71 Attendance*		E/C Difference*	t Ratio
		E	C		
Selmer	4.06	3.86	3.500	+.36	0.109
Athens	7.52	3.64	6.78	-3.14	-0.926
Wichita	6.30	5.25	7.07	-1.82	-1.401
Dallas	10.25	1.29	7.72	-6.43	-6.467
Rockland	5.65	6.89	5.10	+1.79	2.018
Fresno	6.12	6.16	6.68	-0.52	-0.648
McComb	4.09	6.70	8.83	-2.13	-1.781
Seattle	5.57	5.81	5.41	+0.40	0.483

* All entries are percentage days absent.

TABLE X-1. (Continued)

C. GRADE 3

Grade	Median 69-70 Attendance*	Estimated Mean 70-71 Attendance* E	C	E/C Difference*	t Ratio
Selmer	3.62	4.79	5.94	-1.15	-0.805
Athens	4.68	4.92	5.62	-0.70	-0.493
Wichita	5.18	5.27	5.61	-0.34	-0.219
Dallas	7.53	7.59	6.14	+1.45	1.486
Rockland	5.60	5.45	4.69	+0.76	0.956
Fresno	5.52	7.11	6.42	+0.69	0.786
McCombs	6.16	7.46	5.76	+1.70	1.427
Seattle	4.61	4.45	4.34	+0.11	0.147

D. GRADE 7

Grade	Median 69-70 Attendance*	Estimated Mean 70-71 Attendance* E	C	E/C Difference*	t Ratio
Selmer	3.21	2.83	4.16	-1.33	-1.193
Athens	4.09	5.64	5.35	+0.29	0.225
Dallas	8.70	7.14	10.62	-3.48	-2.222
Rockland	5.44	7.72	6.14	+1.58	1.456
Las Vegas	5.23	5.69	7.07	-1.38	-1.419
Fresno	5.16	10.10	6.93	+3.17	3.094
McCombs	3.92	6.98	6.07	+0.82	1.036
Seattle	5.36	6.37	5.10	+1.27	1.325

* All entries are percentage days absent

TABLE X-1. (Continued)

E. GRADE 8

Grade	Median 69-70 Attendance*	Estimated Mean 70-71 Attendance*		E/C Difference*	t Ratio
		E	C		
Selmer	3.06	4.42	4.45	-0.03	-0.039
Athens	6.17	5.50	8.93	-3.43	-2.030
Wichita	5.14	6.50	8.33	-1.83	-1.009
Dallas	5.97	5.83	14.12	-8.29	-4.405
Rockland	5.65	6.65	5.68	+0.97	0.874
Las Vegas	6.62	8.21	8.79	-0.58	-0.491
Fresno	7.39	10.10	8.98	+1.12	1.006
McComb	3.97	7.35	4.76	+2.68	2.726
Seattle	6.48	5.98	6.18	-0.20	-0.285

* All entries are percentage days absent

TABLE X-1. (Continued)

F. GRADE 9

Grade	Median 69-70 Attendance*	Estimated Means 70-71 Attendance*		E/C Difference*	t Ratio
		E	C		
Selmer	4.00	2.03	4.96	-2.93	-5.431
Athens	7.36	5.84	10.68	-4.84	-2.589
Wichita	5.59	7.94	6.58	+1.36	0.859
Dallas	8.49	8.30	10.14	-1.84	-1.041
Rockland	6.25	7.65	7.02	+0.63	0.399
Las Vegas	7.09	9.33	9.38	-0.05	-0.033
Fresno	9.24	9.83	11.07	-1.24	-0.899
McComb	3.02	8.33	3.10	+5.23	4.235
Seattle	7.04	6.62	6.61	+0.01	0.007

* All entries are percentage days absent